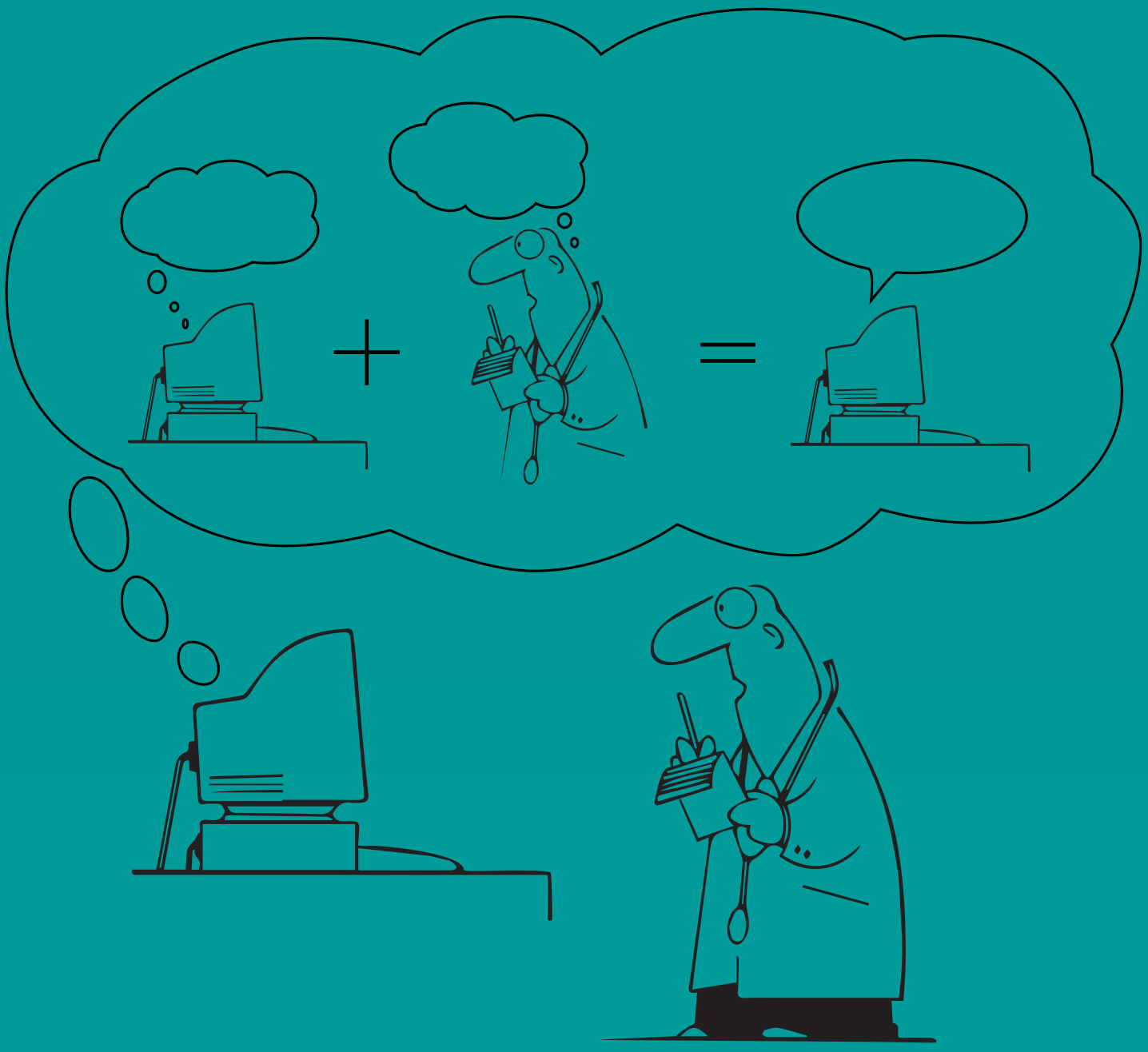

A Cognitive Approach to Modeling Bad News Conversations

Bart van Straalen

A Cognitive Approach to Modeling Bad News Conversations

Bart van Straalen

PhD dissertation committee:

Chairman and Secretary:

Prof. dr. P.M.G. Apers, University of Twente, NL

Promotors:

Prof. dr. A. Nijholt, University of Twente, NL

Prof. dr. D.K.J. Heylen, University of Twente, NL

Assistant-promotor:

Dr. M. Theune, University of Twente, NL

Members:

Prof. dr. F.M.G. de Jong, University of Twente, NL

Prof. dr. G.J. Westerhof, University of Twente, NL

Prof. dr. S.C. Marsella, Northeastern University, USA

Prof. dr. H.C. Bunt, University of Tilburg, NL

Prof. dr. J-J.Ch. Meyer, University of Utrecht, NL

Paranymphs:

Mark ter Maat, PhD

Saskia ter Maat–Akkersdijk, MSc

Human Media Interaction group

The research reported in this dissertation has been carried out at the Human Media Interaction group of the University of Twente.

This book is part of the CTIT Dissertation Series No. 15-354
Center for Telematics and Information Technology (CTIT)
P.O. Box 217, 7500 AE Enschede, The Netherlands. ISSN: 1381-3617.

Game Research for Training and Entertainment

This research has been supported by the GATE project, as part of the work-package 2.2: “Modeling Cognitive Behavior of Virtual Characters”. This project has been funded by the Dutch Organization for Scientific Research (NWO).

This book is part of the SIKS Dissertation Series No. 2015-14
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

ISBN: 978-90-365-3857-2
DOI: http://dx.doi.org/10.3990/1.9789036538572
ISSN: 1381-3617, No. 15-354

A Cognitive Approach to Modeling Bad News Conversations

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. H. Brinksma,
on account of the decision of the graduation committee
to be publicly defended
on Wednesday, May 20, 2015 at 12:45

by

Bart van Straalen

born on December 9, 1980

in Hoorn, The Netherlands

This thesis has been approved by:

Prof. dr. ir. Anton Nijholt, University of Twente, NL (promotor)
Prof. dr. Dirk Heylen, University of Twente, NL (promotor)
Dr. M. Theune, University of Twente, NL (assistant-promotor)

ISBN: 978-90-365-3857-2

Acknowledgements

"I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I —
I took the one less traveled by,
And that has made all the difference."

– Robert Frost, *The Road Not Taken*

Hier, aan het begin van het proefschrift zijn we dan aangekomen bij het eind van de reis. Na een lange weg is het doel bereikt en heeft het de vorm aangenomen van dit proefschrift. En hoewel de weg niet altijd even gemakkelijk is geweest, heeft de manier waarop ik het doel bereikt heb, net zoals in het citaat hierboven, al het verschil gemaakt. In het geval van mijn promotie heeft "the road less traveled" meerdere betekenissen. Aan de ene kant illustreert het het onderwerp van mijn onderzoek, dat zich voor een groot gedeelte heeft gericht op abstracte cognitieve processen en kenmerken. Toen ik in 2007 begon, was dit niet een onderwerp wat erg op de voorgrond stond bij HMI. Het was interessant om te zien dat dit onderwerp in de loop van mijn AIO-schap steeds vaker naar voren kwam.

Aan de andere kant zegt het iets over het verloop van mijn onderzoek en het schrijven van deze dissertatie. Het is niet altijd een gemakkelijk proces geweest. Mijn onderzoek kenmerkte zich regelmatig door een gebrek aan structuur en duidelijke doelstellingen. Daarnaast zorgde het onderwerp ervoor dat gedachtegangen vaak afdwaalde in een filosofische richting in plaats van een praktische / toepasbare richting. Deze dingen hebben er voor gezorgd dat ook het schrijven van het proefschrift een proces was dat veel tijd en energie heeft gekost. Het heeft bij tijd en wijle een grote mate van volharding en doorzetting vereist om het proefschrift af te maken. Het feit dat ik het uiteindelijk helemaal afgerond hebt, is hetgeen dat voor mij al het verschil heeft gemaakt.

Natuurlijk heeft het nooit zo ver kunnen komen zonder de hulp van de hulp en steun van een groot aantal mensen. Ik wil graag beginnen met mijn promotoren en assistent-promotor te bedanken. Anton, je hebt als mijn promotor altijd grote interesse getoond in mijn onderzoek en het onderwerp ervan, en zorgde je ervoor dat ik jaarlijks haalbare doelen stelde om de voortgang van het onderzoek te garanderen. Heel erg bedankt daarvoor. Als tweede wil ik Dirk bedanken. Zonder jouw hulp had er waarschijnlijk nog minder structuur in mijn onderzoek gezeten. Niet alleen

heb je mij begeleid in het opbouwen van mijn onderzoek, maar hebben we ook vaak gediscusieerd over de inhoud ervan. Hoewel het regelmatig te abstract en theoretisch naar jouw mening was, kon je me toch vaak in de goede richting sturen en mij helpen het geheel tastbaarder te maken. Mijn dank daarvoor. Ik hoop dat je tevreden bent over het uiteindelijke resultaat. Verder wil ik ook graag Mariët bedanken. Hoewel je pas later in mijn onderzoek bent aangesteld als assistent-promotor, heeft het een wereld van verschil gemaakt. Met name jouw reviews van de hoofdstukken en je commentaar hebben ervoor gezorgd dat het proefschrift beter is geworden dan dat ik het ooit alleen had kunnen doen.

In addition I would like to thank the members of my committee, for taking the time and effort to read and comment on this thesis. I hope you all have enjoyed reading it and were inspired by it to find out more about human, cognitive processes and how they can be incorporated into computer systems.

Also, a big 'thank you' to the people involved in the GATE-project and especially those working on the same work-package as I. John-Jules, as work-package leader you were always keen to know how my research was progressing and what I thought about the topic of its contents. Also thanks to the other major members of the work-package: Michael, Maaike and Karel. Although we did meet often, you often provided me with new perspectives to my research.

Wanneer ik zeg dat dit proefschrift niet geschreven had kunnen worden zonder de hulp van anderen, heeft dat vooral betrekking op mijn beide paranimfen. Mark, niet alleen ben ik blij dat ik jou tot mijn goede vrienden mag rekenen, maar ook als collega was je een grote hulp. Het feit dat we ongeveer rond dezelfde tijd zijn begonnen bij HMI en dat onze onderzoeken veel met elkaar gemeen hadden, zorgde ervoor dat we van gedachten konden wisselen over praktische en theoretische concepten. Ondanks het feit dat je er onchristelijke werktijden op nahield (wie werkt er nou van 08:00 tot 16:00?) heb ik veel profijt gehad van onze regelmatige gesprekken en ook van je betere programeer-skills.

Mijn andere paranimf, Saskia, is mijn held met betrekking tot het verzorgen van de opzet, layout en cover van dit proefschrift. Zonder jouw LaTeXen photoshop kwaliteiten had ik nooit mijn proefschrift kunnen afronden. Je stond altijd klaar om mij te helpen als ik weer eens wat stuk had gemaakt in de layout of er niet meer uitkwam hoe een tabel in elkaar zat. Hiervoor bedankt.

Naast het feit dat jullie mij enorm geholpen hebben met de inhoud en het uiterlijk van het proefschrift, wil ik jullie met name bedanken voor jullie vriendschap en continue interesse in hoe het met mij en mijn proefschrift ging. Dankzij jullie was mijn tijd in Enschede een stuk plezieriger. Ik heb veel genoten van onze avonden samen met grote hoeveelheden thee dan wel wijn en hoop in de toekomst nog vaak leuke tijden met jullie door te brengen.

Uiteraard wil ik ook graag al mijn collega's bij HMI bedanken voor een gezellige werksfeer. Met name dank aan mijn kamergenoten uit het AIO-aquarium. Van de oude garde, Ivo, Wim, Thijs en Herwin, tot de nieuwe groep, Robbie en Merijn, en de tijdelijke toevoegingen zoals Hanna en Andrea. Dankzij jullie heb ik erg genoten van mijn periode in Enschede. Met veel plezier heb ik met jullie gediscussieerd over verschillende aspecten van mijn onderzoek. Ook heb ik genoten van de leuke activiteiten

samen, zoals guitar-hero in de experiment-ruimte (nee beveiliging, we zijn echt bezig met onderzoek), een virtual character overrijden met een viruele pickup truck, het lezen van piraten-verhalen (no really, it's science!) en het collectief ontwijken van Wim's uitnodigingen om mee te gaan naar onderwater-hockey.

Verder wil ik ook Dennis bedanken. Door het stellen van kritische vragen heb je mij vaak laten nadenken over wat het nu precies was wat ik wilde vertellen en of dat wel correct was. Dank aan Ronald, voor de grappen en de onderbrekingen van het dagelijkse ritme, maar ook voor bieden van een alternatief perspectief waarop er naar mijn onderzoek gekeken kon worden. Charlotte en Alice, grote dank voor jullie hulp bij alle organisatorische en administratieve zaken.

Veel dank gaat ook uit naar al mijn vrienden. Helaas kan ik niet iedereen hier persoonlijk noemen, maar allemaal heel erg bedankt voor jullie steun, interesse en het verzorgen voor afleidingen. Een paar mensen wil ik even in het bijzonder noemen. Sander, Parcival, Machiel en Barry. Naast jullie interesse in hoe het vorderde met mijn proefschrift, verstrekte jullie mij ook regelmatig met de nodige ontspanning. Dankzij jullie kon ik regelmatig de realiteit even achter me laten en mijzelf verplaatsen in een wereld waar de meeste problemen opgelost kunnen worden door er met een bijl hard op te slaan. Grote dank daarvoor. Jeroen en Andrea, bedankt voor de gezellige dagen in Utrecht of Hoorn, dat er nog vele mogen volgen. Bas, bedankt voor de broodnodige alcoholische versnaperingen op zijn tijd. En Paul: na vier jaar onderzoek naar emotie-theorieën weet ik het nu eindelijk zeker: Kou is géén emotie!

Verder ook grote dank aan mijn familie. Mijn ouders, zonder wiens steun en vertrouwen ik dit niet al die jaren had kunnen volhouden, pap en mam: Bedankt! Daarnaast ook mijn zus en haar gezin, wie naast hun interesse ook voor de nodige afleiding hebben gezorgd.

En op de laatste maar tegelijk ook op de eerste plaats wil ik mijn vriendin José bedanken voor haar steun, geduld en liefde. Lieve José, het was niet altijd gemakkelijk om vol te blijven houden en mijn dissertatie af te maken, maar jouw toewijding en steun waren daarbij een grote hulp. Een deel van dit proefschrift is daarom ook jouw verdienste en daarvoor wil ik je enorm bedanken. Het is ontzettend fijn om te weten dat jij in mij blijft geloven ook al doe ik dat zelf soms even niet. Ik lief jou heel veel!

Contents

Part I

Literature study and Data collection

1

Chapter 1: Introduction

The idea of building a machine with which one can engage in a conversation has a long tradition. Soon after the advent of the digital computer, researchers in Artificial Intelligence, or computational linguistics more precisely, established a research field that investigates the possibilities of turning the digital computer into a machine that you can talk to. The field of spoken dialogue systems has produced some useful applications in which spoken or written language is the mode of interaction , but it seems that we are far removed from building a machine that - based on its conversational skills - can be mistaken for a human being (Turing, 1950). There are quite a few challenges that need to be faced.

It is difficult to make a computer to generate and understand human language. There are several different reasons why this is. First of all, it is not trivial to automatically recognize the words that are being uttered. Speech recognition has come to a stage where useful applications such as dictation systems have been developed, so there is some progress here. However, human language contains many words and phrases that are ambiguous, and deciding on which meaning was intended is not even trivial for humans. Automatic recognition of words in utterances and assigning them meaning is a topic that is addressed in the field of speech recognition (e.g. Jurafsky and Martin (2000))

Secondly, not only does a computer have difficulty understanding what words mean, it also fails to understand what the speaker is trying to achieve (i.e. what he wants) by using the words. Recognizing what the intention of the speaker is, requires more than the ability to correctly interpret the user's words. A classic example that illustrates this problem is the answer "It is raining." to the question "Shall we go out for a walk?".

Thirdly, when humans communicate via spoken language, their speech is often accompanied by non-verbal signals such as facial displays of emotions or gestures of the hands. For a proper understanding of what is happening, the computer also needs to figure out what the different facial expressions and gestures mean. This is important because the nonverbal expressions may provide important clues to the intention behind the words.

In general, understanding words and phrases does not suffice as language use is embedded in a more general context of interaction. This can be illustrated by a simple utterance such as "Put that there." How do we humans know what 'that' and 'there' refers to? This constitutes a fourth challenge. The computer needs to have an understanding of the context in which the conversational behavior is placed to correctly understand the speaker. Certain words and non-verbal signals can have different meanings in different contexts. For example, crying can be an appropriate response behavior to both receiving good news and bad news. Another example is the gesture of a raising a hand with the palm facing outwards. In one situation, this gesture expresses a greeting, while in another situation is might be a signal for 'stop' or 'halt'. These differences in contexts are sometimes difficult to describe in ways the computer can understand them.

However, these difficulties have not deterred researchers from trying to develop dialogue models from which dialogue systems (i.e. computer systems that interact with a user via natural language) can be constructed. Dialogue models aim to represent specific parts of dialogues in such a way that computers are able to understand and hold a natural conversation with a user. A dialogue model can focus on representing intermediate level aspects of dialogues, such as turn taking mechanisms ((Sacks et al., 1974), (Thórisson, 2002), (ter Maat, 2011)), or it can aim to represent higher level aspects, such as the cognitive handling or selecting of conversational behaviors, making it more a cognitive model than a dialogue model (for examples, see ACT-R and SOAR).

In order to construct a dialogue system that is able to hold natural conversations that involve complex topics, not only does the underlying dialogue model need to be extensive, but also the communicating interface has to be sufficiently expressive. For complex conversations (e g. emotion-filled conversations) it is not sufficient to have a simple text-based or even speech-based interface to represent the systems conversational behaviors, because, in such conversations, non-verbal signals and displays play a large role in bringing across the meaning of the conversational behavior and the context in which it is performed. In these cases, Embodied Conversational Agents (ECAs) are a more appropriate form of dialogue system.

ECAs are dialogue systems that are able to interact with the environment via a physical or (more often) virtual representation of a human or human-like body. Such a body is capable of performing verbal as well as non-verbal conversational behaviors. The last few decades have produced a score of ECAs. For examples see: Rea (Cassell et al., 1999), Max (Kopp et al., 2003), GRETA (Poggi et al., 2005) and Elckerlyc (van Welbergen et al., 2010). Although some ECA systems have been constructed as an academic exercise to examine the possibilities of artificial conversational partners, most ECAs focus on performing a specific task. Such tasks include, but are not limited to, information providing tasks (e.g. NUMACK (Kopp et al., 2005) and Gandalf (Thórisson, 1997)), coaching tasks (e.g. exercise counseling agent (Bickmore and Sidner, 2006)) and tutoring tasks (e.g. STEVE (Rickel and Johnson, 1997) and INES (Hospers et al., 2003)).

The dialogue models in task-focused systems often contain a strict protocol or set of rules that indicates which conversational behaviors the ECA needs to perform to

fulfill the task and the sequence in which these behaviors need to be performed. The sequence of an ECA's conversational behaviors is often pre-set so that the course of the conversation corresponds with the progression of the associated task. For example, the protocol in a dialogue system that focuses on tutoring might state that the ECA first needs to perform conversational behaviors that explain the basic steps of the study-material, before it can perform conversational behaviors the address the more advanced steps.

However, natural conversations very often do not follow such strict sets of rules. The selection of appropriate conversational behaviors by the participants is not predetermined by such a fixed protocol. Instead, it is the result of taking into account a participant's current mental state and the conversational behaviors performed by his interlocutor. In this thesis we use the term *interlocutor* to indicate one of the participants of the conversation, other than the one that is being discussed at that moment. In other words, if the behavior mechanisms of the dialogue agent are being described, we use the term *interlocutor* to indicate the human participant and vice versa.

For example, a participant desires to bring about a certain situation (e.g. an action performed by the interlocutor or a particular configuration of the interlocutor's mental state). Based on the interlocutor's conversational behaviors, the participant has made assumptions about the internal state of the interlocutor. The participant uses these new thoughts and combines them with his own thoughts to determine which response behavior is most appropriate to achieve the desired situation.

In order to construct an ECA dialogue system that performs conversational behaviors in a human-like and adaptive manner, the integrated dialogue model also needs to take these two aspects (i.e. the representation of the mental state implemented in the dialogue system and the assumptions about the internal state of the interlocutor it has made based on the interlocutor's conversational behaviors) into account. Consequently, the dialogue system will be less rigid in its selection of conversational behaviors, more reactive to the user's conversational behaviors and more proactive in pursuing the goals it has been programmed to fulfill.

The term *mental state* (also called the *internal state*) is used in this thesis to indicate the collection of an individual's mental faculties that shape the person's cognitive processes and bring about and shape his or her behaviors. These mental faculties (also called *mental features* or *internal state features*) include, amongst other things, the person's thoughts, desires, feelings and motivations. (Chapter [4] will discuss these features in more detail.)

1.1 Research questions

The aim of this thesis is to provide more insight into the mental processes involved in the selection of conversational behaviors that can be performed in a natural dialogue. Having a better understanding of these processes allows us to construct dialogue systems that act in a more human-like manner than current dialogue systems. A dialogue system that acts in a human-like manner characterizes itself by the way it selects and performs its conversational behaviors and how its components and rules function. Using representations of human cognitive processes and features to construct the ar-

chitecture of a dialogue system (and the dialogue model upon which it is based) is a strong and straightforward way of trying to get a dialogue system to act in a similar way as humans do. When comparable processes and features are used to process and select conversational behaviors, it is more likely that the outcome (i.e. the conversational behaviors performed by the dialogue system) will also be comparable to human behavior.

We believe that including human-like mental processes and features could allow such a dialogue system to hold complex conversations in a more natural manner than other dialogue systems. For example, many current dialogue systems do not take the previous, present and future mental states of the user into account when selecting an appropriate conversational behavior. More specifically, they do not form any expectations about how their behavior may affect a user mentally. Consequently, these dialogue systems are unable to actively select conversational behaviors that alter the user's mental state in a specifically intended way. As a result, they are less capable of directing the course of the conversation towards completing their conversational goals.

Especially the manner in which the conversational behaviors might affect the user's emotions and his social relation with the dialogue system is left unattended. This is fine when the dialogue systems need to perform simple, practical tasks such as explaining how to operate a machine (Rickel and Johnson, 1997), giving directions to a certain location (Theune et al., 2007) or making reservations (Traum, 1993). Even in dialogue systems that need to perform complex mental tasks which are guided by strict protocols, such as crisis management (Heuvelink et al., 2009) or military negotiation (Gratch and Marsella, 2001), the conversational behaviors that aim to affect the user's emotions and his perception of the social status are often underrepresented. Instead the conversational behaviors in such dialogue systems focus rather on completing the rule-bound task than on affecting the mental state of the user.

However, when a dialogue system is lacking the capability of taking the user's mental state into account, it fails to act as a considerate and socially apt human-like counterpart. In such cases, the behavior selection processes need to be adjusted or extended in order to make sure the most appropriate behaviors are selected. In order to determine how these processes need to be adjusted or extended, we pose the following research questions:

1. *What determines the meaning and purpose of a conversational behavior?*
2. *Which aspects of a conversation (within a particular domain) need to be taken into account when determining what an appropriate response behavior is?*
3. *Which cognitive processes and features are involved in processing conversational behaviors and selecting an appropriate response behavior?*
4. *How can such cognitive processes and features be represented in a cognitive dialogue model?*
5. *How does the process of selecting an appropriate conversational response behavior operate?*

We address these research questions by analyzing various conversational behaviors, studying theories and models describing (parts of) the related cognitive processes

and features, and constructing a cognitive dialogue model. The aim of this dialogue model is to illustrate how the cognitive processes and features are related to each other and to demonstrate how the cognitive processes function. In particular, we aim to show that the thoughts and feelings implemented in an agent (about both his own internal state and the internal state of his interlocutor) are taken into account by the agent's processes when appropriate conversational behaviors are selected.

The term *agent* is used in this thesis to indicate an intelligent computer implementation (i.e. a dialogue system) that can be constructed based on the cognitive dialogue model. The implementation needs to be able to autonomously perceive, process, and select conversational behaviors and to interact with a user through multimodal conversational behaviors. Although no dialogue system (i.e. agent) was constructed during this study, we use the term *agent* in order to be able to discuss interactions between a user and a possible dialogue agent system that is based on the cognitive dialogue model.

The cognitive dialogue model has been constructed after studying existing linguistic and psychological theories, methods and models that focus on various aspects of dialogues and cognitive processing. Based on these studies we decided to use representations of the various *cognitive features* that are ascribed to humans, such as **beliefs, goals** and **intentions**, to describe the *elements* that together form the agent's internal state. In other words, the *elements* that form the dialogue model's internal state are representations of the *cognitive features* that make up a human's internal state. We argue that using representations of human cognitive features in the construction of the dialogue model has several advantages.

First of all, according to the theories and models studied, the cognitive processes involved in handling, selecting and performing conversational behaviors in humans consist of manipulations of the associated internal state features. Because the elements in the agent's internal state are representations of human cognitive features, it stands to reason that the processes in the agent that manipulate these elements also should to be similar to those humans employ. Such human-like processes and behaviors are desirable in situations where the dialogue agent should play the role of a participant of the conversation, for example in systems that focus on tutoring tasks or training simulations.

Secondly, by having human-like internal state elements and performing human-like processing methods, a dialogue agent (for example in a tutoring or training system) is able to explain why it performed a certain behavior, in a way that is easy for users to understand. This is because the elements and mechanisms the agent uses to process and select conversational behaviors are intuitive to grasp for humans, as they have acquired an understanding of these elements (i.e. cognitive features) and mechanisms in dealing with other people during the course of their life.

Thirdly, as the agent knows how its internal state elements are related to the conversational behaviors it performs, it can make assumptions about the relations between the user's conversational behaviors and the user's internal state features. Subsequently, because the components in both internal states (i.e. the agent's elements and the user's cognitive features) and the way they are related to their respective conversational behaviors are comparable, the agent "understands" how its conversational

behaviors might influence the internal state features of the user.

This understanding has an effect on the agent's behavior selection. To fulfill its goals, the agent can direct the conversation in two ways. On the one hand, it can perform its own conversational behaviors (e.g. asking a question) and thus getting its goal fulfilled (e.g. getting / knowing information). On the other hand, the agent can use its conversational behaviors to affect the user's internal state and thereby influence the selection of the user's conversational behaviors. For example, a tutoring agent may have the goal to see the user perform a specific task. During the conversation, the agent may provide the user with bits of information (i.e. which the user can use to form new beliefs) so that the user can independently complete the task.

When holding a conversation (i.e. processing perceived conversational behaviors and selecting and performing response behaviors), a dialogue agent manipulates the elements in its internal state. The rational part of processing and selecting conversational behaviors is done by manipulating the representations of beliefs (including assumptions about the interlocutor's internal state features), goals and intentions. Representations of these three kinds of cognitive features are essential for holding a conversation. However, in order to enable a dialogue agent to act as human-like as possible, the dialogue model upon which it is based needs to be further augmented. This can be done by adding representations of additional kinds of cognitive features to the agent's internal state, such as emotions and the dispositions the agent has with respect to the social relations that exist between itself and its interlocutor. Manipulating the representations of emotions and social dispositions constitutes the more irrational part of processing and selecting conversational behaviors. A large part of the processing and selecting of the conversational behaviors performed in natural and complex conversations is associated with these internal state elements.

The inclusion of elements that represent emotions and social dispositions into the dialogue model may influence the functioning of three mechanisms. Firstly, these elements can affect the agent's processing of perceived conversational behaviors. Secondly, they can affect the selection of an appropriate response behavior and the manner in which that response behavior will be performed. Thirdly, including the concepts of emotions and social relations into the dialogue model enables the dialogue agent to form beliefs about the emotions and social dispositions of its interlocutor.

We propose to derive that the relations between the various internal state features (or elements) and the conversational behaviors can be derived by analyzing the conversational behaviors themselves. To that end, we introduce the term *behavior properties* in chapter [6]. The properties describe how the features of the internal state are displayed or represented by (parts of) the conversational behaviors. Examples of behavior properties include the `displays of emotions` (e.g. crying and smiling, or a raised tone of voice when angry) and the `displays of social dispositions` (e.g. the degree of politeness used in the behaviors). Other behavior properties are the `intended effect` and the `type of the conversational behavior`. For example, the intended effect of the conversational behavior containing the utterance "Is the tumor removed?" is to get the interlocutor to provide the speaker with an answer, in other words to gather information. A complete overview of the behavior properties related to internal state features that are used in this thesis, is provided in section

[6.1].

To identify how the various behavior properties are expressed in actual conversational behaviors we have performed a detailed analysis of a video clip in which a conversation is portrayed. We choose to analyze a particular type of conversations, namely bad news conversations conducted between a physician and a patient. The reason for selecting this type of conversations is that they contain natural and complex behaviors since the topics are often discussed reluctantly and the utterances are frequently prevaricating. Furthermore, the conversational behaviors in bad news conversations contain a wide array of emotional and social displays. The conversation in the video clip contains over 30 conversational behaviors, performed by two different interlocutors. We performed an in-depth analysis that provides us with sufficient information to identify the behavior properties and thus the relations between the conversational behavior and the internal state features.

Using the results of the analysis, we make assumptions about the relations between the behavior properties, and the elements and processes that make up the speakers' internal states. For example, we argue that the intended effect is a property of a conversational behavior through which the speaker's intention can be brought about. Subsequently we argue that the intended effect of a conversational behavior allows us to directly make assumptions about the intention underlying the behavior. Recognizing and understanding the intended effect of a conversational behavior is thus essential for interpreting the intention of the speaker.

Another example is that certain social dispositions towards an interlocutor manifest themselves through the degree of politeness in the speaker's behaviors. In order to strengthen the assumptions we make, a questionnaire concerning the conversational behaviors in the same video was conducted. Via the questionnaire we asked several people what they believed the speakers in the video clip were thinking and feeling while they performed specific conversational behaviors. The aim of the questionnaire was to see whether the assumptions about the relations between the conversational behaviors and the elements and processes of the internal states could be validated through general consent.

Based on the results of the analysis and the questionnaire, we performed a categorization of the conversational behaviors performed in the video clip. The categorization groups the conversational behaviors together based on similarities between the properties of the behaviors. The behaviors are placed into a category if their properties match the values of the variables that were chosen for each category. More information about the categorization process is presented in chapter [5].

In addition to organizing the conversational behaviors, the categorization is also used by our cognitive dialogue model in two ways. On the one hand, the dialogue model's cognitive processes use the information about the properties of behavior (e.g. assumptions about the relations between them and the speaker's internal state features) that is expressed through the categories to create a new internal state. On the other hand, the dialogue model selects an appropriate response behavior from one of the categories. These two ways are discussed in more detail in chapter [6].

1.2 Thesis outline

The thesis is structured in the following way. In **chapter [1]** we outlined the topic of this thesis. We presented the motivations for performing this specific research and defined our research questions. In addition, we presented the approaches we have taken to come to satisfying answers to these research questions.

In **chapter [2]** we discuss several approaches to study language and conversations that are useful in the construction of a dialogue model. We identify the subtle difference between sentences and utterances, discuss the meaning of language used in a conversation and show how to distinguish between different types of meaning. In order to be able to relate conversational behaviors to internal state elements and processes, speech act theory and dialogue act theories are also studied and discussed, along with the conclusions we draw from these theories. Furthermore, we discuss various types of dialogue models to show why we use an agent-based approach for our own dialogue model.

Chapter [3] covers the domain of bad news conversations. In order to gain a better understanding of medical bad news conversations we describe several protocols on how such conversations should be held, as well as various theories and models that explain how people deal with bad news situations, both from the perspective of the bringer of bad news and the perspective of the recipient. Furthermore, we give an overview of the most common difficulties that may occur when conducting a bad news conversation and show how people handle these difficulties.

In **chapter [4]** we discuss several theories and models that describe possible approaches to representing the features and processes of a person's cognitive, internal state that are involved in processing conversational behaviors. We focus on those theories and models that contain similar features and processes as those we aim to use in the construction of our own cognitive dialogue model.

First, we discuss several theories that describe how the rational part of an agent's internal state, i.e. its beliefs, goals and intentions, might be represented. Next, we discuss several theories and models that focus on processes that explain how conversational behaviors can cause emotions (appraisal) and how emotions can influence the selection of appropriate response behaviors (coping). We discuss how the various internal state features (i.e. beliefs, goals, intentions and emotions) can be represented abstractly and how the appraisal and coping processes are connected to the elements of the rational part of the internal state. We also briefly discuss several theories concerning the social relations that play a role in a bad news conversation and how such social relations can influence the processes involved in the construction or managing of the other internal state features.

In **chapter [5]** we present an analysis of the conversational behaviors performed in a simulated bad news conversation. In addition, we present the findings of an online questionnaire we conducted and the analysis we have done on these findings.

In **chapter [6]** we present a categorization of the conversational behaviors performed in the aforementioned simulated bad news conversation. The categorization specifies several properties of each conversational behavior used in the conversation, such as a behavior's content and the manner in which it is performed. We specify a

list of those properties that people use in the processing of conversational behavior. We illustrate how these behavior properties are related to the various types of internal state features represented in our cognitive dialogue model. We explain how the elements of the dialogue model's internal state can be constructed or altered.

In addition, we compare the theories and methods presented earlier in this thesis and the various processes and representations of internal state features contained in our own cognitive dialogue model. Furthermore, we compare our dialogue model to dialogue models that are similar in approach and set up, and indicate were the differences lie.

In **chapter [7]** a pen-and-paper example of the workings of our dialogue model is presented. The entire process from perceiving a conversational behavior, to interpreting and processing that behavior, to forming or selecting new internal state features, to selecting an appropriate response behavior is presented for several conversational behaviors. In addition, we present the preliminary work we have done on constructing an implementation of our cognitive dialogue model and connecting it to various visualization components.

Chapter [8] presents the conclusions we have formed. We also specify some of the benefits and shortcomings of our cognitive dialogue model and of the gathered data. Conclusively, we provide several suggestions on how our research and our dialogue model might be extended through future work.

2

Chapter 2: Deconstructing dialogues

Every day people use language, either in written text or spoken out loud, seemingly without difficulty. But when examined closely and studied vigorously it becomes apparent that the use of language consists of very complex processes. Within the field of linguistics all aspects of language and how language is used are studied.

Specific studies of language form include, but are not limited to, morphology, phonology, prosody and syntax. Studies of the meaning of language are concerned with topics such as semantics (i.e. how meaning is inferred from words), and pragmatics and sociolinguistics (i.e. how meaning is inferred from the relationship between sentences and the situations in which they are used). Other studies focus on the use of language in relation to other fields of research and include topics such as anthropological linguistics, psycholinguistics and neurolinguistics.

In this thesis, particular interest is placed on spoken dialogue and how the use of language in a conversation influences the conversational behaviors of the interlocutors. In addition, we focus on the question of how the performed conversational behavior is related to the thoughts and feelings of the speaker as well as those of the listener.

This chapter discusses various methods and theories that describe the act of performing conversational behaviors, i.e. how language is used, and also what these conversational behaviors are composed of. Prominent in this type of approach is Austin's theory of locutionary, illocutionary and perlocutionary acts (Austin, 1962), which has formed the basis for many other theories that have been constructed over the years. We start in section [2.1] by giving definitions of the linguistic units we will be discussing in this thesis. Section [2.2] discusses the term *meaning* with respect to the use of language. In section [2.3] we discuss Austin's theory in more detail and present Searle's adaptation of this approach (Searle, 1969), (Searle, 1975). Section [2.4] describes a more recent approach to describing conversational behaviors, namely through dialogue acts. All these methods and theories provide us with knowledge on how to define different aspects of language. To show how these aspects can be connected in a dialogue model we describe several prominent dialogue models in section [2.5]. Finally, section [2.6] presents our conclusions.

2.1 Sentences & Utterances

When discussing language in the context of spoken dialogue, the linguistic units that are used by the speakers are called utterances rather than sentences. Although the terms *sentence* and *utterance* are used to indicate different things, they are closely related to each other. The term *sentence* is most commonly used to refer to abstract linguistic units that correspond to the highest level of the grammatical system that structures language. Bloomfield defines it as "an independent linguistic form, not included by virtue of any grammatical construction in any larger linguistic form." ((Bloomfield, 1933) as cited by (Goodwin (1981), p.7).

Utterances are more difficult to define; there is no general agreement regarding the definition of an utterance. The definition given by Goodwin states that the term *utterance* refers to "the stream of speech actually produced by a speaker in a conversation." (Goodwin (1981), p.7). This stream of speech includes "the entire vocal production of the speaker - that is, not only those sounds which could be placed in correspondence with elements of sentences, but also phenomena such as midword plosives, inbreaths, laughter, crying,"uh's", and pauses." (Goodwin (1981), p.7). Utterances are defined by Clark as follows: "Utterances are the actions of producing words, sentences and other things on particular occasions by particular speakers for particular purposes." (Clark, 1996). This relation between sentences and utterances is also stated by Lyons, who states that "as a grammatical unit, the sentence is an abstract entity in terms of which the linguist accounts for the distributional relations holding within utterances. In this sense of the term, utterances never consist of sentences, but of one or more segments of speech (or written text) which can be put into correspondence with the sentences generated by the grammar." ((Lyons, 1969) as cited by (Goodwin (1981), p.7)). Whereas Goodwin seems to restrict the term *utterance* to spoken language, Lyons also makes reference to written language.

Clark's definition of utterances indicates that utterances are placed in a specific context, determined by occasions, speakers and/or purposes. This contrasts with the notion of the term *sentence* as "sentences are . . . abstracted away from any occasion on which they might be used, stripped of all relation to particular speakers, listeners, times and places." (Clark (1996), p.128).

Although the definitions about the nature and meaning of utterances presented in the previous paragraph differ slightly from each other on several points, they also present several interesting aspects of utterances that assist us in using the term practically. For example, both Goodwin and Clark indicate that the context in which the term utterance is used is formed around an active performance or action on the part of the speaker. In addition, the definitions show us that utterances consist of segments of speech (or written text) that relate to linguistic units in the grammatical structure. These linguistic units can be a single word, a part of a sentence, a whole sentence or even a sequence of sentences. Furthermore, another characterizing aspect that marks the difference in status between utterances and sentences, particularly when the utterances are expressed through speech, is that utterances can contain various kinds of disfluencies. Examples of such disfluencies may include the following things. A speaker may start to produce an utterance, change his mind and start over. Alterna-

tively, he may make a mistake during his performance and may decide to repair the mistake. Or, if the speaker gets the impression the listener is not understanding him, the speaker may interrupt himself and provide an explanation. In addition to the presence or absence of disfluencies, vocal performances of an utterance are characterized by prosody, which assists the speaker to get his meaning across to the listener. This auditive component of an utterance can also convey additional information about what the speaker wishes to achieve by performing his utterance.

A final point to note in the definitions is that utterances need not only involve the production of linguistic units such as words and sentences, but, as Goodwin points out in his definition, also include the production of non-linguistic units such as laughter, etc.

In the next section, we discuss how the term meaning is used in relation to the use of language.

2.2 Meaning

The term *meaning* is very ambiguous and a proper explanation is needed to indicate what is meant by this term. A review of the literature shows that meaning can be attributed to a wide variety of things, each of which are often given a different designation. We will use the term *expression* in this thesis, to indicate that to which we attribute a particular meaning. The term *expression* includes, but is not limited to, words, certain combinations of words (e.g. "the month of May"), sentences and signals.

The term *signal* is used to indicate particular kinds of expressions. Clark defines signals as "deliberate human acts" (Clark, 1996). Using this definition and the definition we formed in section [2.1], we subsequently equate signals with utterances when signals are viewed in the context of language use. This inference is corroborated by Clark: "So when I use the terms *utterances, speakers* and *speaker's meaning*, I normally intend *signals, signalers* and *signaler's meaning*." (Clark (1996), p.128).

A distinction can be made between two types of meaning, namely the meaning of a signal itself (i.e. *signal meaning*) and the meaning that can be attributed to the thoughts of the speaker that underlie signals (i.e. *speaker's meaning*). Because both terms use the word *meaning*, it is important to understand the difference between them. The difference between these two terms is much clearer in non-English languages, where these concepts are known by different names. For example, in Dutch the term *betekenis* is used to indicate signal meaning and *bedoeling* to indicate the speaker's meaning. In German, the terms *Bedeutung* and *Gemeintes* are used respectively.

The term *signal meaning* is used to express the way in which a signal should be interpreted. This is what is meant when we refer to the term *meaning* in the philosophical sense of the word, i.e. with a certain 'reference' and with a certain 'sense' (Frege, 1892). The *reference* of an expression is the object or event to which the expression refers, while the *sense* of an expression is the manner in which the object or event is referred to by the expression i.e. its mode of presentation. It is through its mode of presentation that an expression conveys particular knowledge. In his ar-

ticle, Frege explains the division of meaning into sense and reference by using the expressions “the evening star” and “the morning star” as an example. Both expressions describe the celestial body Venus, but the first expression is used when Venus is observed during the evening while the second expression is used when Venus is seen during the morning. Now, the reference of both expressions is the same (i.e. the object the planet Venus) but the sense of the expressions is different, as the expression “the morning star” conveys different properties of the object it is referring to than the expression “the evening star”. This shows that when the expression “the morning star” is used as the mode of presentation, something different is meant than when the expression “the evening star” is used, even though both expressions refer to the same object.

The term *speaker’s meaning* is used to indicate what the speaker is trying to communicate by performing the signal. In other words, it is the intention of the speaker that is conveyed through the utterance to the listener, i.e. the reason why the utterance was performed by the speaker. In order to get the desired response from the listener, it is important that the speaker’s intention is conveyed through his utterance and that the listener receives and understands the speaker’s meaning. The question is how we can recognize or extrapolate the speaker’s meaning from the utterances the speaker performs. For this, a more thorough description of utterances and how they are used is needed. In the next section we discuss several approaches on how utterances can be described and how these descriptions relate to the meaning of the speaker and of the signal.

2.3 Speech acts

As can be seen in the definitions presented in section [2.1], one of the important aspects of an utterance is that by making an utterance, the speaker is performing a kind of action in a conversation. Thus, an utterance can be seen as an action. However, the use of the word *action* in the context of speaking is rather ambiguous. In his paper, Austin illustrates this with the following example: “we may contrast men of words with men of action, we may say they **did** nothing, only talked or **said** things; yet again, we may contrast only **thinking** something with actually **saying it** (out loud), in which context **saying** is **doing** something.” (Austin (1962), p.92). The idea that utterances can be regarded as actions is strengthened by Austin by introducing the notion of performatives to contrast constatives. Constatives are utterances that only assert or state something that can be judged to be true or false, for example saying “the color of the ball is red.” Performatives, on the other hand, are utterances such as the following.

- I order the doctor to tell the truth.
- I address the reader of this thesis.
- I promise you everything will be alright.

These alter the state of the world simply by being uttered by a speaker, just like other actions change the state of the world by being performed. According to Austin, performatives, such as making a promise or giving an order, are actions. Constatives,

on the other hand, are described as sayings, for example making a statement or giving a description. This suggests that, because utterances are actions (according to the definitions given above), utterances can only be performatives and not constatives. However, this is clearly not the case, and Austin continues to talk about the distinction between performative utterances and constative utterances. The fact that utterances are both constatives and performatives can only hold if we treat the act of uttering an utterance as also performing an action.

This notion is strengthened by Searle (1989), who states that uttering a statement and uttering a description are just as much actions as promising and ordering. This raises the question how to distinguish between the different types of utterances and how they relate to actions.

According to Austin it is expedient to go back to the fundamentals of the use of language. The basis of Austin's theory is that all utterances are actions that consist of uttering a word or sentence to get listeners to recognize what the speaker means. Based on Austin's theory, Searle argued that every utterance can be described by means of three specific kinds of acts, that all hold at the same time. He dubbed these specific kinds of acts *speech acts*: According to Austin it is expedient to go back to the fundamentals of the use of language. The basis of Austin's theory is that all utterances are actions that consist of uttering a word or sentence to get listeners to recognize what the speaker means. Based on Austin's theory, Searle argued that every utterance can be described by means of three specific kinds of acts, that all hold at the same time. He dubbed these specific kinds of acts *speech acts*:

- *Locutionary act*: The act of uttering an expression which has a particular meaning.
- *Illocutionary act*: The act of getting the listener to recognize the speaker's meaning by uttering an expression.
- *Perlocutionary act*: The act of causing the listener to produce certain consequential effects upon his thoughts, feelings or actions by uttering an expression. What these effects are, is based on the listener's understanding of the meaning of the expression.

The term *locutionary act* is used to indicate the act of 'saying something'. Locutionary acts bear no relation to the listener or to the underlying motivation of the speaker, but only convey the action of performing an utterance. Consequently, locutionary acts only contain the meaning of the signals (i.e. the utterances), but not the meaning of the speaker.

On the other hand, the term *illocutionary act* is used to describe just that: the speaker's meaning. An illocutionary act is that part of performing the utterance that aims to get the listener to recognize what the underlying intentions of the speaker, i.e. the speaker's meanings, are. What has caused the speaker to perform this utterance and to what end? Note that the illocutionary act does not describe whether or not the speaker's meaning is received correctly by the listener or that his intention is properly understood, only that this is the purpose of the utterance. While the utterance is directed at the listener and the illocutionary act is performed to get the listener to recognize the speaker's intention, the effect the utterance has on the listener is not contained in the term *illocutionary act*.

It is obvious that one cannot perform a locutionary act without also performing an illocutionary act at the same time, as every utterance that has a signal meaning is also performed with a specific purpose in mind and thus also contains a speaker's meaning. The intentions that the speaker tries to convey through his utterances, are called *illocutionary forces* in Austin's theory. Examples of illocutionary forces provided by Austin include, but are not restricted to:

- asking or answering a question,
- giving information, a warning or an assurance,
- announcing a verdict or an intention,
- pronouncing a sentence, (i.e. performing a locutionary act)
- making an appointment, an appeal or a criticism,
- making an identification or giving a description.

This list of illocutionary forces gives a good indication what kind of intentions can possibly underlie an utterance, but it is by no means complete or structured.

In an attempt to deal with this lack of structure and completion Searle composed an ordered categorization of illocutionary forces into which all speech acts could be classified (Searle, 1969), (Searle, 1975). According to Searle, illocutionary acts can be categorized via what he calls their *illocutionary points*. Illocutionary points are part of the illocutionary forces and they describe the "publicly intended perlocutionary effect" of a speech act (Clark (1996), p. 134). The publicly intended perlocutionary effects are the actual purposes of the speech act, while the rest of the illocutionary force consists of particular presuppositions. These presuppositions are certain background beliefs related to an utterance that are mutually known or assumed by the speaker and the listener, so that the utterance can be considered appropriate for the context. For example, the utterance "please close the window", presupposes, among other things, the beliefs that there is a window and that it is not closed already. The purposes, i.e. illocutionary points, of speech acts (and thus of illocutionary acts) may be to get the listener to do something, or to make the speaker commit to doing things. In all, Searle created five categories into which speech acts can be divided, based on their illocutionary points:

- **Assertives**: By uttering an assertive the speaker is committing himself to the truth of the expressed proposition. Assertives are characterized by the following illocutionary point. The speaker tries to get the listener to form or attend to the belief that the speaker is committed to a certain belief himself. Examples of illocutionary forces that are associated with assertives are stating, suggesting, boasting, concluding, swearing, and denying.
- **Directives**: Directives are speech acts that are uttered by the speaker in an attempt to cause the listener to take a particular action. How forceful directives are depends on the type of illocutionary force expressed by the speaker, ranging from mild suggestions to stern commands. The following list are examples of illocutionary forces that are related to directive illocutionary acts: asking, begging, inviting, ordering and commanding. Directives can further be divided into two major classes: requests for actions (often expressed via commands and

suggestions) and requests for information (often expressed via questions). Note that an attempt to cause the listener to take the action of forming a belief is an assertive and not a directive.

- **Commissives**: The illocutionary acts that fall in the commissives category are characterized by the speaker's expression of his commitment to some future course of action. Such commitments can be expressed through illocutionary forces such as promising, offering, vowing, betting and predicting.
- **Expressives**: Expressives are used when the speaker wants to express his attitude with respect to some state of affairs that concerns him or the listener. Most of the time these are expressions of the speaker's emotions, but they also include feelings that are formed by social conformity. Illocutionary forces that are associated with expressive illocutionary acts include: greeting, thanking, welcoming, apologizing, congratulating and condoling.
- **Declarations**: The illocutionary acts associated with utterances that bring about a change in the state of the world in accord with the proposition of the utterance are called declarations. Declarations are often expressed via performatives, which were mentioned in the beginning of this section. Illocutionary forces that are associated with the declarations are: naming, pronouncing, resigning, defining and obviously declaring. While declarations often work by virtue of conventions of institutions such as the law (a judge pronouncing a penalty or sentence), the church (a minister blesses somebody) or a company (boss promotes somebody), this is not strictly necessary. Anyone can make declarations or definitions without being bound by such institutions.

Utterances can possibly fall into several categories and it is the job of the listener to figure out what the speaker's meaning of the speech act was. Searle called the process of understanding the speaker's meaning of the utterance through the recognition of the illocutionary act, the *illocutionary effect* of that utterance. The recognition of a speaker's meaning is of great importance in order to have proper communication. However, not only recognizing and understanding the meaning of the speaker by the listener is important, but that the listener produces an appropriate response is equally crucial.

Utterances can also be described through a third kind of act, namely the *perlocutionary act*. A perlocutionary act is an act that produces certain effects on the thoughts, feelings or actions of the listener on the basis of his interpretation of what the speaker means. If, for example, the speaker utters the sentence "Your disease cannot be cured", the listener might - not unlikely - form the belief that his disease is incurable. The effects that are caused in the listener's thoughts, feelings or actions by the speaker's utterance are called the *perlocutionary effects* or *perlocutions*. It is important to note that the perlocutionary effects might differ from what the intended effect of the speaker's utterance was, i.e. the illocutionary point of the utterance. If the listener experiences sadness and shock - and even disbelieve - as a consequence of the speaker's utterance, the production of these feelings is a perlocutionary effect, even though it was not intended by the speaker. Furthermore, perlocutions are not necessarily caused by the listener's understanding of the signal meaning or the speaker's meaning that are part of the utterance. If the listener is addressed in an unknown

foreign language, certain consequential effects will be produced, even if he has no idea about the meaning of the utterance.

2.4 Dialogue Act theories

Even though the term speech act comprises all three types of acts, it is generally used to describe illocutionary acts rather than the other two types. In more recent work, researchers have taken Searle's categorization of speech acts, i.e. illocutionary acts, as a basis and expanded this notion, modeling more types of the intentions that underlie speakers utterances. While these theories are expansions on speech acts, a variety of terms is used to indicate this concept. Terms other than speech acts that are used throughout the years include *communicative acts* (Allwood (1976), Sadek (1991)), *conversational acts* (Traum and Hinkelman, 1992), *conversational moves* (Carletta et al., 1997) and *dialogue acts* (Bunt, 1994). The term *dialogue act* can perhaps be seen as the most generic when discussing the use of language in dialogues (Traum, 2000). In the next section, a more in depth overview of the theory of dialogue acts is presented.

2.4.1 Dialogue Acts

The term dialogue act can be defined as follows: "A dialogue act is a unit in the semantic description of communicative behavior produced by a speaker and directed at a listener, specifying how the behavior is intended to change the information state of the listener through the listener's understanding of the behavior." (Bunt, 2005) Dialogue acts are comparable with speech acts and any of the other terms mentioned above, in that they are concepts that are used to analyze and describe the meaning (both the signal meaning and the speaker's meaning) of utterances that are performed in a dialogue. However, the concepts used in the dialogue act approach are considered to be more formal than the traditional concepts used in speech act theory. The dialogue acts concepts have a well-defined formal semantics, which, according to Bunt, is fundamental for constructing a structure in which conversational behavior can be described. It is important to note that dialogue acts are used to analyze and describe the interpretations of conversational behavior by an observer (which can include the listener) rather than the behavior itself. Dialogue act theory focuses on how the conversational behaviors are related to the internal states that hold for the speaker as well as the listener. Bunt expresses this rather nicely in his paper: "To say that a speaker performs a certain type of dialogue act is to say that he produces an utterance (possibly linguistic, or gestural, or multimodal) of which the analysis of its meaning involves an intended type of change of (the internal) state / context which can be described by the communicative function and the semantic content of that dialogue act." (Bunt, 2005).

Similar to Austin's approach to speech acts, Bunt distinguishes three aspects of dialogue acts that describe how an utterance can be interpreted. These aspects are the *utterance form*, the *semantic content* and the *communicative function*. The *utterance form* determines the manner in which an utterance is performed and may include prosodic properties (when speech is used), layout information such as italics or capital

letters (in case of written text) and the use of punctuation or pauses. It indicates the utterance's mode of presentation. The *semantic content* of an utterance is the information that the speaker makes available to the listener. This information states what the dialogue act is about: which objects, events, situations, substances, etc. does it refer to? What propositions involving these elements are considered, using what properties, relations ...? The semantic content of a dialogue act can be compared to the description a *locutionary act* provides. It corresponds to the *signal meaning* of the performed utterance. The *communicative function* of a dialogue act expresses what the listener is supposed to do with the semantic content, i.e. it describes how the semantic content is to be used by the listener to update his thoughts and feelings. In other words, the communicative function expresses the purpose of the dialogue act and thus the `intended effect` of the utterance, i.e. the speaker's meaning. The communicative function thus corresponds to the same things as the *illocutionary forces* in speech act theory or to be more precise, the *illocutionary points*.

In addition to the more formal semantics of dialogue acts with respect to the description of utterances, there is another important difference between the dialogue act approach and speech act theory. A major problem with speech act theory is that it only allows for utterances to be described by a single speech act, while an utterance might simultaneously express multiple purposes (Allwood, 2000). For example, the utterance "I will be there around eight o'clock" in response to a question "What time will you be there" has the following purposes: The listener *confirms* that he understands the speaker (i.e. the question-poser), in addition he *informs* the speaker that he (i.e. the listener) will be there at eight and finally the listener *promises to perform* a certain action (i.e. to be there at eight). Dialogue act approaches allow for these different purposes of a single utterance to be labeled by multiple dialogue act types (Allen and Core (1997), Bunt (2005)).

With the rapid developments of computer-aided systems and technologies in the last few decades, dialogue acts have been used for more purposes than what speech acts originally were designed for. Currently, dialogue acts are used for the following endeavors (Bunt, 2005):

- To support conceptual analysis of natural, human dialogue
- As building blocks in the interpretation and generation of utterances in dialogue systems. This development can assist us in the construction of complicated and more extensive dialogue systems.
- To annotate (a corpus of) dialogues, both in human-human interactions as in human-computer interactions.

In order to be able to achieve these endeavors, the semantics of dialogue acts need to be structured. This structuring has resulted in the development of several taxonomies that formally categorize the different types of dialogue acts and provide each category with a clear and formal description. In a sense, these taxonomies are a continuation and extension of the categorizations Austin and Searle developed for the different types of speech acts. Two of the major taxonomies that have been developed in the last two decades are DAMSL **(Dialogue Act Markup in Several Layers)** (Allen and Core, 1997) and DIT++ (Bunt and Black, 2000) and (Bunt, 2009). Although

originally designed for dialogue annotation, the dialogue act types that are described in both taxonomies can also be used for the other endeavors listed above. In the following subsection we describe the structure and the various dialogue act types that are expressed in the DIT++ taxonomy. We argue that the communicative functions expressed in the DIT++ taxonomy are strongly related to the intentions of the speaker and thus play a large role in the selection process of conversational behaviors. Also, DIT++ takes DAMSL as a basis for its taxonomy and extends it.

2.4.2 The DIT++ Taxonomy

The DIT++ taxonomy is a comprehensive system of dialogue act types obtained by extending the taxonomy of Dynamic Interpretation Theory **(DIT)**, originally developed for information dialogues (Bunt, 1994), with a number of dialogue act types from DAMSL (Allen and Core, 1997). Dynamic Interpretation Theory makes a distinction between two types of goals that utterances in a dialogue can express. On the one hand, a speaker performs certain utterances in order to achieve an underlying internal goal, such as getting particular information or reaching a decision. To achieve such a goal, the speaker needs to do more than just communicate with an interlocutor. It also requires the interpretation and evaluation of new information, the handling of conflicting interests, changing the listener's perception of the situation etc. These types of goals are called *non-communicative goals*. The other type of goals an utterance can express are *communicative goals*. Communicative goals deal with managing the communication itself. They deal with aspects such as ensuring contact, turn taking, monitoring attention, repairing communicative failures, interpersonal relations etc. The communicative goals are to make sure that the conversation runs smoothly, so that it might facilitate the achievement of the non-communicative goals.

In order to achieve communicative and non-communicative goals, respectively *dialogue control acts* and *task-oriented dialogue acts* are used. These two types of dialogue acts are expressed at the highest level of categorization in the structure of dialogue acts. All other dialogue act types are subsumed by these two top level categories.

The communicative function of dialogue control acts is to structure the course of the conversation and may include things such as ***turn-taking, elicitation of feedback*** from the listener or ***self-correction***. The communicative functions of task-oriented dialogue acts may vary widely, depending on the underlying task. For example, in information-transferring dialogues (containing dialogue acts that have the functions to provide and gain information) one often finds e.g. ***questions, answers*** and ***requests***. In negotiation dialogues one often finds ***offers, accepts*** and ***refusals*** whereas in a tutoring dialogue one commonly finds ***explanations, corrections*** and ***verifications***. Concisely, the communicative function of task-oriented dialogue acts is to perform steps in the specified task.

The distinction between dialogue control acts and task-oriented dialogue acts is further characterized by the effect they have on different parts of the context of an interaction. The notion of context of an interaction has a wide variety of connotations in the field of linguistics. Bunt states that the common denominator of these various descriptions of context is that "they all refer to factors relevant to the understanding of

communicative behavior." (Bunt, 1994). Within the Dynamic Interpretation Theory, Bunt make a distinction between five kinds of context:

- **Linguistic context:** This contains the properties of the expressed linguistic material (textual or spoken). These properties include things such as prosodic information (in the case of spoken utterances) and layout information and punctuation (when written text is used).
- **Semantic context:** This contains the objects, properties and relations that are relevant to the underlying task, including the progression with respect to achieving the non-communicative goal.
- **Physical context:** This contains the physical circumstances in which the dialogue takes place, the time and place, the communicative channels that are used (e.g. telephone versus face-to-face), the interpersonal distance between the interlocutors (i.e. proxemics), presence or absence of third parties etc.
- **Social context:** This contains the type of interactive situation (e.g. a formal conversation or chat with friend) and the roles of the interlocutors, as characterized in terms of the communicative rights and obligations that are contained in their interpersonal relation.
- **Cognitive context:** This contains the elements which the internal state of the interlocutors is comprised of. This includes the interlocutors' beliefs, goals, intentions and emotions as well as their states of processing that are related to the perception, interpretation, evaluation, production and execution of conversational behavior.

The linguistic context is determined by the utterance form aspect of dialogue acts performed, regardless of whether these dialogue acts are task-oriented or dialogue control acts. This follows from the fact that the utterance form describes how the conversational behaviors are expressed. In other words, changes to the linguistic context of an interaction can be made by task-oriented dialogue acts or dialogue control acts. Similarly, both task-oriented dialogue acts and dialogue control acts can influence the make-up of the cognitive context of a conversation. However, in this case changes to the context are caused by the communicative functions of the dialogue acts, rather than by the utterance form.

The difference between the two types of dialogue acts manifests itself in the fact that task-oriented dialogue acts also cause changes in the semantic context in addition to the linguistic and cognitive contexts, while dialogue control acts additionally affect the physical and social contexts. Similar to changing the cognitive context, changing the semantic, physical and social contexts is caused by the communicative functions of the dialogue acts.

It is thus the purpose (i.e. the communicative function) of a conversational behavior that determines how the interlocutors perceive the conversational behavior, not the type of dialogue act. Consequently, the DIT++ taxonomy shows a categorization of similar communicative functions, rather than a categorization of particular

dialogue act types. This is similar to the manner in which Searle constructed his categorization of speech acts, which was based on the illocutionary points of the speech acts.

The DIT++ taxonomy groups together communicative functions that are "characterized by notions of intuitive conceptual similarities." (Bunt, 2009). For example, the communicative functions of dialogue acts concerned with ***grabbing, keeping, giving*** or ***accepting*** the sender role are called **Turn management functions**. Communicative functions of dialogue acts that contain ***greeting*** and ***thanking*** behaviors fall into the group **Social obligations management functions**. The groups in which the communicative functions can be placed are referred to as *dimensions* ((Allen and Core, 1997), (Bunt, 2009)). These dimensions have strict boundaries which helps to structure the dialogue act descriptions of utterances performed in a conversation. However, not all communicative functions can be categorized neatly. For example, the communicative function of a dialogue act that has ***question*** as a type, does not fit into one specific dimension. The question can be asked about something that is task-related, but it can also be about closing a specific subtopic in the conversation or about whose turn it is, to contribute to the conversation. So it can be said that ***questions*** belong to all these dimensions. The same holds for communicative functions of dialogue act types such as ***answers, statements, requests, offers, explanations*** etc. The DIT++ taxonomy clusters the communicative functions that are broader than one dimension and structures them separately from communicative functions that can be placed in a single dimension. Consequently, the taxonomy consist of two parts:

- a set of clusters of General-purpose functions,
- a set of clusters of Dimension-specific functions.

The set of **General-Purpose functions** consist of two broad categories, both of which are divided further into two subcategories. The first category, **Information Transfer functions**, consists of those functions whose aim it is to obtain or to supply information. Its subcategories are **Information-Seeking functions** (for obtaining information) and **Information-Providing functions** (for supplying information).

Information-Seeking functions include different kinds of ***questions***, i.e. ***direct, indirect, yes-no questions, wh-questions***, etc. **Information-Providing functions** include dialogue act types such as ***informs, agreements, elaborations***, but also ***answers, confirmations*** and ***disconfirmations***.

The second category of **General-Purpose functions** consist of **Action Discussion functions**. These functions aim to introduce certain actions into the conversation that may or should be performed by either the speaker, the listener or by both. **Action Discussion functions** are further divided into **Commissives**, when the action is directed towards the speaker, and **Directives**, in the case the speaker is putting pressure on the listener to perform a particular action. Examples of **Commissives** are: ***offers, promises, bets*** and ***predictions***, whilst **Directives** include ***instructions, requests, suggestions, orders*** and ***commands***.

Contrary to General-purpose functions, **Dimension-specific functions** are grouped into clearly defined dimensions. The DIT++ taxonomy specifies ten distinct dimensions. Nine of the ten dimensions contain dimension-specific communicative functions that deal with creating and maintaining the conditions that lead to a successful

conversation, i.e. communicative functions that handle the management of dialogues. The communicative functions that only deal with handling the management of the dialogue are called **Dialogue control functions**. The tenth dimension, **Task / Activity**, contains dimension-specific communicative functions that express particular domain-related purposes of a dialogue act.

Table [2.1] is presented in (Bunt, 2006) and shows the ten dimensions specified in the DIT++ taxonomy, examples of Dimension-specific communicative functions and typical expressions per dimension. The DIT++ taxonomy has a more extensive set of dimension-specific functions than the one presented in Table [2.1], but Table [2.1] allows for a clear image to be formed. A complete overview of the taxonomy is included in appendix [A] and can also be found at the following website: http://dit.utv.nl [1].

Dimension	***Dimension-specific Communicative functions***	***Typical expressions***
Task/Activity	OpenMeeting, CloseMeeting, Appoint, Hire, Fire	domain-specific fixed expressions
Auto-Feedback	PerceptionNegative EvaluationPositive OverallPositive	*Huh?* *True.* *OK.*
Allo-Feedback	InterpretationNegative EvaluationElicitation	*THIS Thursday.* *OK?*
Turn Management	TurnKeeping TurnGrabbing TurnGiving	final intonation rise hold gesture with hand *Yes.*
Time Management	Stalling	slowing down speech fillers
Contact Management	ContactChecking	*Hello?*
Own Communication Management	SelfCorrection	*I mean . . .*
Partner Communication Management	PartnerCompletion	completion of partner utterance
Discourse Structure Management	DialogueActAnnouncement TopicShiftAnnouncement	*Question.* *Something else.*
Social Obligations Management	Apology Greeting Thanking	*I'm sorry.* *Hello!, Good morning.* *Thanks.*

Table 2.1: Examples of dimension-specific communicative functions

Most of the names of the dimensions quite clearly express what the communicative functions contained in them do with respect to the course of the conversation, only **Auto-Feedback** and **Allo-Feedback** might require some clarification. The communicative function of the dialogue acts grouped in the **Auto-Feedback** dimension is to provide feedback to the utterance of the interlocutor, i.e. A performs a conversational response behavior as feedback to something B said. For example, B says "It's better to be well, than to be sick.", to which A responds "That's true.".

Allo-Feedback dialogue acts are about A's beliefs about B's processing of A's last utterance. That is, whether A believes that B has correctly understood and processed A's

[1] Accessed on 24-03-2014

conversational behavior. Some of these allo-feedback dialogue acts focus on eliciting feedback from B. For example, A says "The appointment is at seven o'clock." immediately followed by the utterance "Right?". This second utterance is an Allo-Feedback dialogue act.

The review of the DIT++ taxonomy provides a good insight into which kinds of communicative functions a dialogue act can have. As communicative functions represent the purposes of the dialogue acts (i.e. their `intended effects`), determining the communicative functions of conversational behaviors provides us with a basis for making assumptions about the underlying intentions of the speaker. The descriptions of some of the communicative functions described in the DIT++ taxonomy also express information about the speaker's beliefs that underlie the dialogue act associated with the communicative function. For example:

- If interlocutor A (the speaker) performs an ***inform*** type of dialogue act aimed at interlocutor B (the listener), A wants (i.e. has the intention) to make information p known to B (i.e. change B's beliefs about p). Information p forms the semantic content of the ***inform*** act. Furthermore, A assumes (i.e. believes) that information p is correct.
- If B performs a dialogue act in return that has an ***agreement*** communicative function, B (now the speaker) believes that A (now the listener) believes the semantic content of the previous ***inform*** act (i.e. p) to be true. Also B has adopted the new belief that p is true and wants to convey this to A.

Although the communicative functions provide us with ample information, not all features from a speaker's internal state that underlie his dialogue acts are clearly expressed in the DIT++ taxonomy. And while more recent work by (amongst others) Bunt (e.g. (Bunt, 2011) and (Bunt et al., 2012)) focuses on mechanisms that describe how the communicative function and the semantic content of a conversational behavior might be related to the speaker's internal state, it doesn't always specify clearly what the exact features that form these internal states are and how they should be defined. The study of these mechanisms is certainly an interesting direction of research. Unfortunately, due to time constraints we were unable to study this particular work in detail. We believe that a more extensive study of the DIT++ taxonomy might prove beneficial for improving our cognitive dialogue model and is something that should be taken into account in future research.

2.5 Dialogue models

Dialogues do not only consist of an exchange of utterances. They are also characterized by processes that structure a conversation on a meta-level. In order to represent the processes involved in a conversation, a dialogue model can be constructed. The manners in which dialogues can be modeled are quite diverse as can be the reasons for modeling them. For example, in order to represent dialogues in a theoretical manner, a formal logic approach can be used. Another possible manner to model dialogues is through an empirical approach, which uses psychological experiments concerning

conversations as a basis to construct representations of the involved processes. Alternatively, one can take a practical approach to dialogues, wherein the focus of the model is not to represent the processes in the most realistic manner, but to make practical decisions on how to construct a good working dialogue system.

In this section we describe different types of dialogue models that can be found in the literature. The dialogue models that are described focus on the management of dialogues rather than on the parsing of utterances performed in the dialogue. We distinguish four main approaches that can be used to model dialogue management, i.e. finite state-based and frame-based approaches, information state approaches and agent-based approaches.

2.5.1 Finite state automata

One manner in which dialogues have been modeled frequently in the past is through finite state automata. For this type of modeling, a conversation is perceived as a sequence of steps which are performed by the interlocutors. Each step represents a state in which the dialogue can be. Consequently, a finite state automaton consists of a finite set of predetermined states covering all dialogue acts that might be performed. Transitions from one state to another occur when a certain event or condition is triggered. Each state can have several follow-up states, allowing for a variety of alternative paths through the dialogue. More information about the theory of the finite state-based approach can be found in (Cohen, 1997).

Types of conversations that are typically modeled through finite state automata include reservations, booking of tickets or any other type of conversation where the number of possible follow-up states is limited for each conversation state. This is because these types of conversations are often well-structured but not very extensive. Due to the limited number of possible states the conversation can be in, it is quite easy to draw up a model for the entire conversation. In addition, the required vocabulary and grammar for each state can be specified in advance. However, this tends to make a system which uses a finite state model very inflexible and rigid.

One of the limitations of dialogue systems that use a finite state approach is that they cannot infer extra information provided by the user and thus proactively moderate the conversation. They can only follow the predetermined steps specified in the system. The event or condition that triggers the transition to the next state will be the only piece of information from the users utterance that is stored by the system. This can lead to unnatural conversations:

System (state 1): “Welcome. Where do you want to travel?”
User: “I want to travel from Liverpool to London this Wednesday.”
(The system stores End-location: London. This triggers a transition to state 2)
System (state 2): “Where do you wish to depart from?”
User: “I told you. I want to take the train from Liverpool this Wednesday.”
(The system stores Start-location: Liverpool. This triggers a transition to state 3)
ystem (state 3): “When do you want to travel?”

Another limitation is that the finite state automata approach suffers from problems

with flexibility. This is because making a change in the possible course of the conversation means that numerous, if not all, states in the model need to be redesigned.

2.5.2 Frame based approach

The problem of inflexibility is dealt with in the frame-based approach to human-machine dialogues. Instead of modeling dialogues according to a sequence of predetermined states, information about the state of the conversation (or task) is gathered in frames consisting of forms or slots. The slots in the frames are filled in based on the dialogue acts of the human interlocutor. The dialogue system uses these frames to determine which utterance should be used to respond to the interlocutor. Take for example a train timetable information system that uses a frame-based dialogue model. If a user says to the system "I want to travel from Liverpool to London this Wednesday", the system can fill in the slots of the following frame:

```
[Travel:
    Origin = Liverpool
    Destination = London
    Date = next Wednesday
    Time = .
]
```

The system now only has to inquire after the time the user wants to travel. If, in a single dialogue turn, the person only told the system that he wants to travel to London, the system needs to select conversational behaviors to inquire after the date, the time and the place of origin. The system would use the same frame to check which information it already received and which information it still needs to acquire.

For a finite state automaton with the same task, each configuration of the values of the four attributes would need to be integrated into the system. This means one transition for "I want to travel from Liverpool to London this Wednesday", another transition for "I want to travel to London", yet another transition for "I want to travel this Wednesday." etc. For each added attribute, the number of dialogue states needed increases exponentially. A system using the frame-based approach is much easier to extend when extra attributes must be added and is more flexible with respect to processing received information.

In addition to the benefit of increased flexibility, the frame-based approach also processes information a user is providing that has not been asked for. For example, if the system asks "Where do you want to travel?" and the person replies with "I want to travel to London, on Wednesday", the system can also use the information "on Wednesday" to fill another one of the slots in the frame. This allows the system to ask more directed questions to get the information needed to fill in the slots that are still open. Furthermore, it allows the system to, proactively (i.e. on its own), insert information in its conversational behavior and filter out obsolete questions. An example using the train timetable information system again:

System: "Welcome. Where do you want to travel?"
User: "I want to travel from Liverpool to London this Wednesday"

System: "You want to travel from Liverpool to London this Wednesday."
System: "At what time do you wish to depart?"

However, the frame-based approach still only caters for simple and direct types of conversations and it also relies for a great deal on the user to provide information. For a good overview of dialogue systems using a frame-based approach see (Bui, 2006).

2.5.3 Information state approach

An alternative to finite state and frame-based approaches is an information state approach. The information state approach has been developed to overcome the limitations of finite state approaches and frame-based approaches. In the information state approach, one central information state represents the state of the dialogue. All input flows into this state and the selection of response utterances is based on its content. New data can be stored and processed quite easily, regardless of whether this information has an external source (such as information provided by the user) or and internal source (information already stored in the internal state). Also, because the system has continuous access to all its information it can produce robust responses to unexpected inputs. According to Traum and Larsson (Traum and Larsson, 2003) an information state approach consists of five components:

- The elements which are used in the information state, such as beliefs, intentions, conversational structure, etc.
- The formal representation of these elements.
- A set of updating and formation rules that keep the information state up to date.
- A set of dialogue moves that will trigger the updating of the information state.
- An update strategy that determines which update rules are applied at a given point.

The problem with the information state approach is that specifying the formation and updating rules is much more complex and time-consuming than in finite state and frame-based approaches. One way of dealing with this problem is by using existing data to learn how to cope with the input. This can be done for example via Markov Decision processes or Reinforcement learning techniques.

2.5.4 Agent based approach

Building on the information state approach are agent-based dialogue models. Agent-based approaches can handle conversations of greater complexity than the previously explained approaches. The basis of agent-based dialogue modeling is that conversations can be viewed as a collaborative process between two intelligent interlocutors, both of which have goals and intentions they wish to see fulfilled. If this is the case, a dialogue model can be constructed that is goal directed. This means that a system using this type of model is not only reactive, but also proactively can form a plan to achieve its goal. In addition to having goals and intentions, the system can use an interaction history, domain knowledge, context knowledge and a user model to reason about the intentions of the interlocutor as well as to infer these intentions from

the interlocutor's utterances and possibly to make assumptions about the goals and beliefs of the interlocutor. As these inferences are more general than explicit dialogue states or forms in frames, a system with an agent-based approach can deal with new or unexpected input in a much more robust manner. The problem is that this type of approach is difficult to model as the processing of the input not only influences a single state or elements in the information state, but also has an effect on the system's own goals and intentions. However, as we are attempting to model dialogues that are complex in nature and where the internal state elements of both interlocutors are taken into account when selecting conversational behaviors, an agent based approach seems most suitable.

2.6 Conclusions

In this chapter we have discussed different manners in which utterances in conversations can be described, ranging from what utterances exactly are to how and why they are used in a conversation. Understanding the meaning of the language used in conversational behaviors helps us to make assumptions about the relation between the behaviors and the elements of the internal state that hold while performing the behaviors. Parts of Austin's and Searle's theories on speech acts suggest how the meaning of conversational behaviors can be determined. Specifically, the illocutionary act expresses what the speaker wants to achieve with his conversational behavior, i.e. what his intentions are. Based on this notion, together with the research presented in section [2.4] where we discuss dialogue acts, we argue that there is a distinction between an intended effect of a conversational behavior (called the communicative function in dialogue act theories) and an intention of a speaker. An `intended effect` is a `property of a conversational behavior` that expresses what that conversational behavior is aimed to achieve. An **intention** is an feature in the speaker's internal state that represents a situation that the speaker wants to be the case. This distinction is used in our dialogue model (presented in chapter [6]) to relate a property of a conversational behavior (the `intended effect`) to an assumed feature (i.e. element) of the speaker's internal state; (the **intention**). In addition, the communicative functions of the conversational behaviors allow us to make assumptions about other internal state features, such as the speaker's **beliefs** and **desires / goals**. More information about internal state features can be found in chapter [4].

In order to enable the dialogue model to efficiently process the observed conversational behaviors, the behaviors are grouped into specified categories. The categories in the dialogue model's behavior-base are based on the dimensions in the DIT++ taxonomy (discussed in section [2.4.2]). The different categories and the process of categorization is explained in further detail in section [6.1.1].

By applying the theories and methods described in this chapter, we can determine what the meaning and the purposes of a conversational behavior are. Based on this information we can start to make assumptions about the different internal state features that are associated with the conversational behavior. Not only can we make assumptions about the features that hold in the internal state of the speaker, we also argue that the conversational behaviors affect the internal state features of the

listener, and subsequently influence the selection of his response behaviors. This information is used during the construction of the dialogue model presented in chapter [6].

Our study of different types of dialogue models resulted in our decision to use an agent-based approach to handle the cognitive processes involved in a conversation. This is because an agent-based approach seems best suited to model an intelligent conversational system that has its own goals and intentions. The reason for constructing a dialogue model that includes cognitive features is that we believe that such a model will produce conversational behavior that is natural and human-like. The support for this belief is that by modeling the features and processes as realistically and human-like as possible, the resulting conversational behavior will also be more human-like.

3

Chapter 3: Bad news conversations

In this thesis we look at a specific type of conversations, namely conversations that deal with bad news situations in a medical environment. With the steady increase in life expectancy and the accompanying issues of chronic illness and quality of life, the importance of being able to properly conduct bad news conversations for health care workers also increases as well as the need for understanding how bad news conversations affects both doctors and patients.

One of the reasons why this domain was selected is because bad news conversations contain a wide variety of emotional and social displays. By observing and interpreting these kinds of displays we can make more detailed assumptions about the configuration of the speaker's mental state that has led to the performance of the speaker's conversational behavior. Another reason is that both interlocutors, often a patient and a doctor, need to take the internal state of the other into account in order to perform appropriate kinds of response behaviors so that the conversation can run smoothly.

This chapter discusses the nature of conversations that handle the delivery and reception of bad news. In section [3.1] we explain what the term *bad news* means, by looking at definitions given in the literature. Section [3.2] discusses various issues that can cause difficulties during bad news conversations for both the bringer and the recipient of bad news and presents several options of how such difficulties might be handled. In section [3.3] we reflect on what can be learned from the different methods and techniques discussed in this chapter and how we can incorporate this information in the construction of our cognitive dialogue model, with the focus on the role of emotions and social values in the selection of conversational behaviors.

3.1 Defining bad news

When asked "what is bad news?", most people will have an intuitive notion what such information is about. However, they can possibly find it difficult to describe the term *bad news* in a concise and clear manner, as it can be quite abstract what the term means. For example, the question "what amounts to news being bad?" will arise. A

child being told that he cannot have a candy bar will much less likely be conceived as being bad news than a struggling single mother getting fired from her day job. Here a distinction is made in the severity or degree of the bad news. Not getting a candy bar will likely be considered less bad news than losing one's job. But what about a child getting fired from his summer job or the case in which the mother is not single or struggling? The situation in which the bad news is placed also seems to have a large impact on determining the description of bad news. If someone is told that it will it rain for an afternoon they might not interpret it as being bad if it occurs during a two week holiday but it might be seen as disastrous if it will happen during an outside wedding ceremony.

The term *situation* as it is used here, describes the conditions of all the features of the environment (including all the persons that are involved and their thoughts and feelings) that are relevant to the bad news and the relations that exist between them. Whether the recipient takes all the features of the environment into account when processing bad news or only a selection of them, varies from situation to situation. The situation in which the bad news is given can be described in terms of the different types of context presented in section [2.4.2], which together make up the situation.

At first sight it seems less difficult for people to define bad news when it occurs in a medical situation. Medical bad news is often strongly associated with serious illness or situations that are life-threatening. However, "bad news is difficult to define because participants in a medical interaction appraise information subjectively as good, neutral, or bad, and therefore, almost any news has the potential to be bad" (Ptacek and Eberhardt, 1996). Furthermore, the difficulty of giving a clear definition is not only influenced by the situation that is described by the bad news, but also by the subjective assessment of the receiver. Each person responds differently to hearing bad news, regardless of the situation. Some people might assess certain bad news as being disastrous whilst others might find the same bad news only mildly inconvenient. As a consequence, they will differ in defining whether the news is bad. This shows that defining bad news in a medical situation is no less difficult than defining bad news in any other situation.

It is clear that creating a satisfying definition of bad news is difficult. A review of the literature shows that the majority of the definitions about bad news regarding medical situations display a lot of similarities, both implicitly and explicitly. For example, bad news is defined by (Orlander et al., 2002) as "news that will change a patient's outlook for the future in a very negative way. Such bad news can be about a severe illness, prospect of death or increasing levels of limitations." Fallowfield and Jenkins (2004) define bad news as "any information that produces a negative alteration to a person's expectations about their present and future", while Ptacek and Eberhardt (1996) provide the following definition: "News is bad to the extent that it results in a cognitive, behavioral, or emotional deficit in the person receiving the news that persists for some time after the news is received." Buckman and Kason (1992) provides the following definition: "Bad news is information that has an adverse and serious effect on an individual's view of his or her future, noting that bad news is always a subjective appraisal by the individual receiving the news." This definition is also adopted by other researchers (Baile et al. (2000); Gillotti et al. (2002)).

By taking these definitions into account the following conclusions about bad news can be drawn. Firstly, whether news is perceived as being bad depends on the situation in which the bad news is placed, including the recipient's subjective assessment of the information. The second conclusion is that bad news has a negative influence on various features represented in the recipient's internal state. As we are constructing a dialogue model that is able to produce appropriate and cognitively plausible conversational behavior to be used in a virtual human, we need to try and find out what these features are and how the processes that are related to them operate during a bad news conversation. From the definitions of bad news presented above the following abstract terms relating to elements of the internal state can be obtained: "a patient's outlook", "a person's expectations" and "an individual's view". These terms are somewhat vague and they need to be represented in the dialogue model in a stricter fashion. Also the description of possible results of receiving bad news, i.e. "a cognitive, behavioral or emotional deficit" gives us some insight in how to categorize some of the recipient's mental features. The third conclusion that can be drawn from the definitions is that when receiving bad news, the recipient's thoughts and feelings are influenced for different spans of time. For example, upon receiving bad news the patient can form new beliefs or experience emotions that directly lead to the selection of a response behavior. As a result such beliefs or emotions might be quickly resolved. At the same time, the bad news can cause the patient to form thoughts about possible actions or events that may happen in the future, thereby creating beliefs (i.e. expectations) that last longer.

3.2 Bad news conversations

Holding conversations that deal with medical bad news can be very daunting and complex. The doctor has the task to inform the patient of the bad news situation he is in, but also must be able to steer the patient's thoughts towards accepting the bad news and assisting the patient in making the right decisions for further treatment. Furthermore, he needs to help the patient balance his emotions by offering support, comfort and displays of empathy. The patient does not really have a specific task in the conversation but he tries to cope the bad news as well as he can. Although there is no "best way" to cope with a particular situation, it helps if the patient tries to get as much information as possible.

Even with this clear allocation of tasks, bad news conversations can suffer from various difficulties. These difficulties include all things that interfere with the correct and efficient transference and acceptance of the bad news during the conversation. In subsection [3.2.1], we provide an overview of most common difficulties that are inherent to holding bad news conversations for both the sender as well as the receiver. Subsection [3.2.2] describes several methods that show how these difficulties are dealt with in practice or how theoretically a bad news conversation can be conducted in an efficient and proper manner from the perspective of the doctor. The same is done in subsection [3.2.3] but then from the perspective of the patient.

Analyzing how bad news conversations are held in practice allows us to determine which kind of elements of the internal state are most likely to be involved in hold-

ing bad news conversations and hopefully provide some ideas about the best way to model such elements.

3.2.1 Difficulties while holding bad news conversations

One of the difficulties that arises for the doctor when conducting a bad news conversation is that it is often unclear how a particular piece of bad news will affect a receiver. Buckman and Kason (1992) state that "bad news is always in the eye of the beholder, as one cannot know the extent of the impact bad news has on the recipient." For example, a patient who is told that he has a malignant tumor that is causing pains in his chest, when he was expecting to be told that it was caused by an inflamed muscle is very likely to be shocked. This patient will respond very differently than a patient who has a history with tumors and who will thus have a different attitude towards the situation. In order to make assumptions about the impact the bad news will have on the recipient, the sender of the bad news needs to ascertain the recipient's expectations and understanding of the situation. Hence one of the important axioms of delivering bad news is: "before you tell, ask!" (Baile et al., 2000). By asking questions about how the patient perceives the situation he is in, the doctor is able to form an image of what the patient is thinking and feeling. Subsequently the doctor is better equipped to tailor the manner in which he delivers the bad news to the knowledge and emotions of the patient.

While it is very important for the sender to form assumptions about the thoughts and feelings of the receiver when delivering bad news, it sometimes happens that the sender incorrectly assesses the internal state of the receiver. This is, quite intuitively, a consequence of the fact that the sender does not have direct access to the recipient's thoughts and feelings. The sender's assumptions about the internal state of the receiver are partially formed through his interpretation of the behaviors performed and the expressions displayed by the recipient. Moreover, the sender can make additional assumptions by reflecting on how receiving the bad news would affect his own thoughts and feelings. As mentioned in the previous paragraph, asking the receiver questions about his view on the situation helps the sender to make the assumptions more accurate. When the sender makes assumptions that are incorrect, difficulties and mistakes in the conversation will ensue. For example, before a surgery the doctor explains to the patient that the procedure has a low chance of success. The doctor incorrectly forms the belief that the patient understands the situation. After the surgery it becomes clear that the procedure has failed to deliver the desired results. The doctor will expect that when he informs the patient, the patient will be disappointed but not really surprised. When the doctor delivers this bad news to the patient, the progress of the conversation will be hampered because the patient responds in a shocked manner, which is something the doctor did not expect. Consequently, any mental preparations the doctor had made are no longer appropriate for the conversation.

Another thing that makes it difficult for the doctor to have an efficient bad news conversation is when the patient refuses to acknowledge the bad news. In an attempt to deal with the bad news a receiver might perform behavior in which he denies the bad news situation, for instance by saying the doctor has made a mistake. A second type of response the patient may give is completely ignoring the bad news. Without a

proper acknowledgment from the receiver, the sender cannot be sure that his message has come across to the receiver, which obstructs the flow of the conversation.

Also, being uncertain of the medical prognosis makes it difficult for the doctor to hold an efficient bad news conversation, as are the doctor's feelings of failure to face the patient's needs (Miranda and Brody, 1992). The patient's needs in a bad news situation include a clear picture of what the situation is and sometimes the expectancy that the doctor is going to fix the situation. If the situation cannot be resolved, it can become difficult for the doctor to continue conducting the bad news conversation.

From the perspective of the patient, what makes it difficult to have a proper and efficient conversation is the content of the bad news itself. Receiving bad news almost always has a profound impact on the emotions of the recipient. The high intensities of the emotions cause the selection of conversational behavior of the patient to be determined by his emotions rather than by his rational thoughts. This in turn might hinder the progression of the conversation.

Furthermore, the patient might experience difficulties during the bad news conversation because he does not understand what the doctor is trying to tell him. The cause of this difficulty might be that the doctor uses a lot of medical jargon, but also that the patient ignores parts of information the doctor is giving as result of a coping mechanism.

Regardless of what the difficulties are or who they affect, in order to have a proper and efficient bad news conversation these difficulties need to be dealt with. How this can be done is covered in the next two subsections.

3.2.2 Handling the difficulties – The doctor's perspective

As a response to the difficulties presented in the previous section, various methods and techniques have been designed that enable interlocutors to be better equipped to hold bad news conversations. A review of the literature shows that a majority of the techniques focus on enabling the sender to deliver the bad news in the most efficient and agreeable manner. A good overview of these methods is given in the reviews of the literature on bad news conversations by (Ptacek and Eberhardt (1996) and Rosenbaum et al. (2004)). Also noteworthy is the generalized exposition of techniques for giving bad news that is presented in (Eggly et al., 2006).

One often used technique is the S.P.I.K.E.S. protocol developed by (Baile et al., 2000). The S.P.I.K.E.S. protocol helps physicians to deal with the concerns they experience when having to deliver bad news. Such concerns include uncertainty about the patient's expectations, fear of destroying the patient's hope, fear of their own inadequacy in the face of an incurable disease, not feeling prepared to deal with the patient's anticipated emotional responses and possibly feelings of embarrassment for having previously given the patient an overly positive image of the situation (Baile et al., 2000). The S.P.I.K.E.S. protocol consists of six steps that need to be taken into account when delivering bad news to a patient to deal with these concerns. Each of the letters in the abbreviation indicates one of the steps that need to be taken into account.

The first step in the technique is: **SETTING up the interview**. This step consists of two parts. On the one hand the doctor needs to organize the physical situation for

the conversation. This includes having a location that provides the participants of the conversation with some privacy and means to sit comfortably. If the doctor appears calm and relaxed it is likely the patient will be less stressed. It is also recommended to have one or two family members or significant others present during the conversation to support and comfort the patient. Furthermore, the doctor needs to make certain that there are as few interruptions as possible during the conversation and that he has enough time to have the bad news conversation. Perhaps the most important step in the physical part of setting up the conversation is that the doctor tries to increase the level of rapport with the patient. This can be achieved by maintaining eye contact or by placing a hand on the patient's arm. In addition to preparing the physical arrangement of the conversation, the second part of setting up the interview involves the doctor preparing the conversation for himself. By mentally rehearsing the conversation, i.e. reviewing what the plan is for bringing the bad news and how one can respond to a patient's emotional reactions or difficult questions, the doctor is less likely to falter during the conversation when such situations occur. Note that the doctor's assumptions about the patient's perceptions of the situation and the most likely manner in which he will respond to hearing the bad news is included in this step. However these will not be very strong assumptions as the doctor still has little insight in the patient's perception.

The second step of the protocol is: **Assessing the patient's PERCEPTION**. Here we can clearly see a recurrence of the axiom "before you tell, ask!". In this step the doctor tries to ascertain how the patient perceives his current situation, by asking the patient questions. For example, asking "what have you been told about your medical situation so far" gives the doctor an indication about what the patient is aware of. Analyzing the patient's answers allows the doctor to tailor the plan he made beforehand to better suit the patient's needs and understandings by adjusting the assumptions about the patient's perception. Asking the patient these kinds of questions also helps the doctor to get the patient involved in both the conversation and the bad news situation. Letting the patient express his fears and concerns of the situation before giving him the bad news will often allow him to more readily acknowledge the seriousness of the situation and as a consequence be more accepting towards the bad news.

Step three in the technique is: **Obtaining the patient's INVITATION**. In this step the doctor and patient discuss how much of the information of the diagnosis or the results of tests the patient wishes to hear. By establishing the amount of information and the level of detail the doctor will divulge, the doctor can again adjust his plan for providing the bad news. Also, it is not uncommon for patients to try to reject bad news as the conversation progresses. Shunning or denying information is a proven psychological manner of dealing with bad news. In the event that the patient displays such avoidance behavior, the doctor should try to continue informing the patient about the situation by reminding the patient that they had previously discussed the amount of information the patient wanted to receive.

The essence of a bad news conversation is handled in the fourth step of the S.P.I.K.E.S. protocol, namely: **Giving KNOWLEDGE and information to the patient**. Obviously, it is always difficult to receive bad news. Giving the patient an indication

that the news is not good may cause the shock of receiving bad news to be less severe. Possible ways of indicating that bad news is coming are using phrases such as "I'm sorry to tell you ..." or "Unfortunately, I have some bad news for you." When providing the patient with the specifics of the bad news (i.e. the medical facts), the doctor needs to keep in mind the following things. First, provide the information on the level of comprehension of the patient to prevent misunderstandings. This includes avoiding technical medical terms such as "biopsy" or "metastasis", instead using terms such as "sample of tissue" and "spread of the disease". Secondly, avoid excessive bluntness (e.g. "You have very bad cancer and even if you get treatment you will likely die.") One of the aims of having an efficient bad news conversation should be to make it as stressless as possible for the recipient of the bad news. If the doctor is blunt, the patient is likely to feel isolated during the conversation and angry afterwards which may lead to the patient blaming the doctor for the bad news situation. Consequently this will strongly influence further conversation the doctor and patient may have. The third guideline that should be taken into account by the doctor is to provide the information in small fragments and periodically check if the patient still understands the situation. Often, when receiving bad news the amount of information that is absorbed by the receiver is lessened by the shock of hearing the bad news. Fourth in the case where the prognosis of the patient's medical state is poor, the doctor should avoid using defeatist sentences such as "There is nothing more we can do for you." While it may be the case that the disease can not be cured, there are a lot of other goals that the patient might have at that point, good pain control and symptom relief being two of the most common ones.

After the doctor has delivered the bad news, he reaches step five in the protocol: **Addressing the patient's EMOTIONS with empathic responses**. The patient's reaction to hearing the bad news often consists of two types of responses. On the one hand the patient can give a rational response, such as asking a question to the doctor or commenting on the bad news. Questions are asked by the patient in order to gain more knowledge about the situation or insight in the future, whilst comments can express the patient's thoughts or previous assumptions. In addition to the rational portion of the patient's responses, it is very likely that the patient will display an emotional reaction to the bad news. It has been shown (Buckman and Kason (1992); Ptacek and Eberhardt (1996)) that one of the most difficult challenges of giving bad news is adequately responding to the patient's emotional reactions. The type of emotional response behavior can vary strongly from patient to patient. Typical behaviors range from keeping silent to expressing disbelief or denial about the situation to crying or being angry at the bringer of the bad news, as an expression of shock, isolation and grief. The S.P.I.K.E.S. protocol advises the doctor to make an empathic response in an attempt to offer support, solidarity and to comfort the patient. For making such an empathic response four things need to be taken into account. First, the doctor must look for a sign of emotion in the patient. This may be tearfulness or a look of dismay or shock. Silence on the part of the patient is also an indication. Secondly, the doctor must identify the type of emotion experienced and displayed by the patient and name it to himself. When the emotional display is subtle or when the patient remains silent, the doctor can once again use open questions to inquire what the patient is thinking

and feeling in order to construct a more accurate image about the patient's feelings. The third point is to identify the reasons for the displayed emotion. While this is often directly related to the bad news itself, it helps if the doctor tries to understand the underlying reasons. The fourth point deals with responding to the patient's emotional reactions. First, the doctor needs to give the patient a brief period of time to vent their emotions. Subsequently the doctor makes an expressing statement, in which he lets the patient know that he has connected the displayed emotion with its underlying reasons, thereby indicating his empathy.

An example of all four points is as follows. The doctor gives the patient bad news: "I'm sorry to tell you that the chemotherapy has not decreased the size of the tumor." The patient responds with tearful eyes and says "I hoped the chemo would work." The doctor recognizes the tearful eyes as a display of sadness and disappointment and also identifies the reasons for the emotions: disconfirmation of the hope that the chemotherapy would work causes the disappointment and the consequences of the failed treatment lead to sadness in the patient. Subsequently, the doctor responds by saying: "I know that this is not what you expected. I was also hoping for a better result," and places his hand on the patient's arm. Through this response the doctor indicates to the patient that he has understood why the patient is upset, i.e. is empathic.

All the patient's emotional displays must be adequately dealt with, before the conversation can continue properly. Any unresolved emotions will make the progress of the conversation more difficult, as they interfere with acknowledging and accepting the provided information. By giving empathic responses to patients that display emotions, the doctor helps to calm the patient and shows his support, thereby making it easier for both interlocutors to continue the conversation.

The remainder of the conversation is captured in the sixth step of the protocol: **STRATEGY and summary**. After the patient has had an opportunity to express his feelings, the doctor should steer the conversation in the direction of discussing further treatment of the patient. Having a clear plan for the future will make the patient feel less anxious and uncertain about his situation. However, it is important for the doctor to assess the patient's readiness for receiving such information before continuing. Once again, asking the patient questions to ascertain the situation is the suggested approach. This also provides the doctor insight in the patient's goals at that point, such as symptom control, pain relief or continuity of care. By discussing possible treatment options with the patient the doctor addresses a number of issues. First it gives the patient the idea that the doctor regards his wishes as being important and allows the patient to feel a measure of control. Secondly, the doctor will experience less sense of failure when the treatment is unsuccessful because the responsibility for making the decision was shared with the patient. And thirdly, discussing the future treatment plan together with the patient will make misunderstanding the purpose of the treatment or misconstruing expected results less likely, thereby preventing difficulties in future conversations.

One of the major criticisms of the S.P.I.K.E.S. protocol is that there is no empirical evidence upholding the validity of the technique, i.e. it has not been empirically tested whether the technique is efficient in the context of terminal illness bad news,

whether it offers constructive guidance to physicians (such as reducing stress by removing questions on how to conduct bad news conversations) or whether it offers support to the patients that are the recipients of the bad news (Goldsmith et al., 2008). Goldsmith et al. also state that the technique lacks keying phrases that might help physicians to actually start the bad news conversation or get the conversation back on track when it is spiraling unto the wrong direction. An example of this is when the patient might not understand or denies the seriousness of the situation and subsequently tries to focus the conversation on how to cure his condition, rather than on accepting that the situation is incurable. Another criticism expressed in the study by Goldsmith et al. is, that the linearity of the S.P.I.K.E.S. protocol is sometimes disrupted by the natural flow of the conversation, thereby making it hard to follow through with the stipulated steps.

While the criticisms presented by Goldsmith et al. are specifically aimed towards the S.P.I.K.E.S. protocol for delivering bad news, the criticisms are similar to those applying to bad news giving techniques in general, for example see the studies performed by Eggly et al. (2006).

Although current bad news giving techniques, such as the S.P.I.K.E.S. protocol, generally might provide practical advice and handholds on how bad news conversations should be conducted, they provide little systematic insight into how the doctor's communication skills are actually improved or how these skills influence the communicative behavior of the patient (Gillotti et al., 2002). It can also be questioned whether or not bad news giving techniques have enough theoretical and empirical foundation to be considered valid approaches. Evaluation of the effectiveness of most bad news techniques is often limited to self-reports made by physicians or evaluations of simulated scenarios (Eggly et al., 2006). While bothersome, this lack of empirical data might be partially the result of the difficulty of studying the domain and fact that the techniques are unable to completely reflect the intricacies of medical bad news conversations. The main criticism of Eggly et al. on existing bad news giving techniques is that some of the underlying assumptions are oversimplified and that they do not realistically and sufficiently represent the complexities of bad news conversation. More specifically, the following three assumptions are suggested to be oversimplified.

- **Physicians can plan a bad news conversation.** Whether news is bad depends on the subjective interpretation of the recipient upon hearing the bad news. Ergo bad news conversation cannot be properly planned in advance.
- **Bad news interactions focus on one central piece of information.** The study by Eggly et al. shows that bad news conversations often contain more than one bad news topic.
- **Bad news interactions consist of a physician-patient dyad.** According to the study performed by Eggly et al. patients are often accompanied by one or more companions, who tend to ask more questions than the patients themselves

Assessment of these assumptions and study of videos of real life bad news interactions between doctor and patient (plus possibly any companions) has led Eggly et al. to propose the following modifications to be included in current bad news giving techniques:

- Take the universal precaution of applying strategies for discussing bad news to all interactions in which information is discussed, because any information has the potential to be perceived as bad news.
- Discuss multiple pieces of related and unrelated information in a way that allows patients and their companions to absorb and respond to each piece of information, independently and in context.
- Address the varying needs of all the participants in the interaction, including patients and each of their companions

An additional suggestion Eggly et al. make with respect to the revision of existing techniques is that future analyses of medical interactions and related guideline development should incorporate theoretical perspectives beyond the notion of linearity and causality. This is because observation and analysis of bad news conversations indicates that such interactions are often nonlinear and highly complex in nature. Consequently, trying to conduct bad news conversations with adherence to a strict, linear protocol seems folly. Eggly et al. propose that future technique development may benefit from "a theoretical perspective known as symbolic interactionism, which assumes that actual interactions do not follow scripts: rather interactions emerge out of social relationships, and the interpretation of the interaction (e.g. bad news) emerges from the interaction itself." (Eggly et al. (2006), p. 719).

It is clear that delivering bad news can be a stressful task for physicians and while current techniques might not be completely scientifically validated, they provided physicians with some comfort and handholds to be better equipped to perform these difficult tasks. However, physicians only fulfill one half of a bad news conversation, the other half being fulfilled by the recipients of the bad news: the patients and their companions. How patients and their companions handle the difficulties that might arise in a bad news conversation is discussed in the next section.

3.2.3 Handling the difficulties – The patient's perspective

While studies and overviews on how bad news conversations should best be conducted from a doctor's perspective are quite numerous, there are fewer studies describing how recipients of bad news deal with such information. In order to gain more insight in the workings of bad news conversations, it seems prudent to also study how such news affects the recipients of such information, how they deal with it and how they might respond to hearing bad news. Such studies might yield valuable information that can be used in future bad news delivering techniques as well as providing insight in the processing of information that negatively influences the receiver's mental faculties. Just as there are various methods and techniques that discuss how to best conduct a bad news conversation from the doctor's perspective, different theories have been formulated on how patients handle receiving bad news and how they might deal with difficulties that occur during a bad news conversation. The difference is that the theories discussing the doctor's perspective are often prescriptive, while those discussing the patient's perspective are more descriptive in nature. Theories about patients handling bad news that have been most prevalent so far are Stage Theories of Dying. Other theories include Context of Awareness the-

ory, task-based approaches and Interaction Adaptation Theory. A good overview of theories discussing how to handle bad news is presented by (Copp, 1998).

One of the leading sources of information on how patients deal with medical bad news situations is the Kübler-Ross model (Kübler-Ross, 1969), commonly known as The Five Stages of Grief or The Five Stages of Dying. Being a stage theory, the Kübler-Ross model consists of five discrete mental approaches for dealing with grief and tragedy. Each stage describes behavior that is formed on the basis of the configuration of elements in the internal state of the patient. The stages described by Kübler-Ross result from analysis of conversations with terminally ill patients, in which they described their thoughts and feelings about their situation. While the model is based on conversations with people who were diagnosed with a terminal illness, it has been found that it can also be applied to people that experience grief in other situations, such as the death of a family member or loved one or a child's reactions to parental separation. The Kübler-Ross model arranges the five stages as follows: **Denial, Anger, Bargaining, Depression** and **Acceptance**. Although the five stages are presented here, and elsewhere, in a particular order it is important to keep in mind that the stages are not necessarily experienced by patients in such a linear fashion or even that patients undergo all five stages. It is quite possible for people to experience the stages in a different order, bypass one or several stages, or jump back to previously experienced ones again and again, perhaps even getting stuck in a particular stage. At the same time the stages are not meant to be a complete description of behaviors that are performed when dealing with bad news.

Denial can be described as a conscious or unconscious refusal to accept facts or information relating to the situation the patient is in. The patient holds on to his beliefs that everything is okay and rejects all attempts to alter that perception. In most cases, denial functions as a temporal buffer for the shocking news, allowing the patient to collect himself, process the information at a rate when the patient is ready for it and in time construct more moderate mental manners of dealing with the bad news. If the situation is such that the negative consequences for a large part can be ignored, for example if the patient has no directly detectable symptoms or pains, a patient might run the risk of becoming locked in this stage, unable or unwilling to constructively deal with the situation. In cases where impending death is more obvious such strong denial is often lessened. A dichotomy exists between complete denial, which is often accompanied by expressions such as: "I feel fine, nothing is wrong with me" or "this is not happening to me", and partial denial, where patients often downplay the seriousness of their illness ("all I need is some medicine and I will better in no time") or only accept his situation some of the time. Behavioral responses given by patients in this stage are regularly devoid of the displays of negative emotions one would expect after receiving bad news.

The **Anger** stage of dealing with bad news is characterized by displays of emotions such as resentment and anger. This type of behavior by a patient can be manifested in different ways, depending strongly on who or what the target of the displayed emotions is. Obviously, feelings of resentment, jealousy or envy are aimed towards others, often those around the patient that are unimpaired by illness. The patient might think that what is happening to them is not fair and should happen to someone

else, who they think are more deserving. “Why me?” is one of the most common questions a patient asks in this stage. “Why me and why could it not have been Mr. X?” The patient also may resent his caretakers that rush around him, order unpleasant tests and restrict his movement in the hospital, but at the end of the day can go home and enjoy life. A consequence of this resentment is that the patient’s behavior will be characterized by lots of displays of anger. This displayed anger might be directed towards many different things: towards the doctor, where it might overlap into blame (“Why can t you cure me?”); towards the nursing staff, who either never leave the patient alone or do not check up on him enough; towards other patients, who restrict the amount of care the patient receives by being there; towards companions, who visit too often or not often enough and also towards the patient himself, who is angry with himself for being sick. Another goal of these displays of emotion is to let people know that despite the patient’s limitations and impending death he is still here, he is still alive, lest he gets the feeling he is forgotten as in time he will be, even if this means acting in a manner that will create a negative viewpoint about him. These displays of anger and resentment make it difficult for family and staff to influence and guide the patient’s thoughts and feelings into a state of acceptance. Luckily, in most cases patients often do not linger long in this stage, although exceptions exist.

In contrast with the anger stage, wherein the patient rages and demands answers or a cure, the stage of **Bargaining** takes a different approach. The patient offers an abstract price, such as “good behavior”, to (often) a higher power, which can be a god, the universe or society at large, to get that what he could not get by demanding through bargaining. Often the bargaining is really an attempt to postpone the inevitable; it has a set goal with a set deadline. For example, a patient bargains for “one last Christmas with my family” or “let me live long enough to see my daughter graduate”. Also, bargains include an implicit promise that patient will settle for only this postponement and nothing more. Almost always this implicit promise is broken once the first postponement is achieved. After making it until Christmas, a patient may bargain again to see New Year’s Eve. While the bargains are not always realistically achievable, the patient tries to deal with the situation in a rational way, setting goals and planning ahead, instead of letting his actions be guided by emotions. To some extent bargaining can even be seen as a partial acceptance of the situation. The patient acknowledges that the situation is terminal, but tries to get out of it through a bargain.

In the course of a terminal illness, almost all patients experience the stage of **Depression** when trying to deal with the situation. The depression stage is characterized by strong feelings of emotions like sadness and fear. Expressions such as “I’m going to die soon anyway, so what is the point?” or “Everything is hopeless, why bother with anything?” are quite common in this stage. Within the Kübler-Ross model a distinction is made between two kinds of depression. On the one hand there is **reactive depression**, which causes the feelings to occur as the consequence of the sense of loss a patient is experiencing in the situation. Such loss can take many forms; severely reduced functionality, the loss of an extremity, the sacrifice of life savings to pay for treatments, but also loss of a stable work or family environment. This last loss may even lead to feelings of guilt in the patient. All these senses of loss are direct

consequences of being in a bad news situation such as a terminal illness. However, there is also the function of the patient to prepare himself emotionally for his departure from life. The feelings of grief and sadness that accompany this preparation phase may lead to what is called **preparatory depression** or **preparatory grief**. The biggest distinction between the two kinds of depressions is that reactive depression deals with past or current losses the patient is experiencing and the preparatory depression handles future losses. While this stage is stressful and hard for the patient, it is also an improvement to the stages previously mentioned. Doctors and caretakers are better equipped to deal with the patient without difficulty. Causes for reactive depression can often quite easily be recognized and dealt with, alleviating the emotions that result from them. Dealing with preparatory depression can be more difficult, but displays of sympathy and comforting behaviors such as placing a hand on the patient's shoulder or just silently sitting together often go a long way of helping. Also, the depression stage shows similarities with the bargaining stage in that the patient is already busy with coming to accept the situation.

When the patient has been able to express all his thoughts and feelings and has dealt with them, he will reach a point where he will contemplate his coming demise with a degree of quiet expectation. This stage is known as the **Acceptance** stage. The patient is often at peace with the situation and the caretaker's attention might need to be shifted towards the needs of the patient's family and friends. Often the patient expresses an emotional detachment and objectivity towards the situation and may use utterances such as "It's going to be okay." to indicate he is prepared to die. Interest in his environment diminishes and the patient often wants to be left alone.

Although Kübler-Ross' stage theory of grief is widely accepted in the world, the model has been criticized for a variety of reasons. One of the major criticisms on the Kübler-Ross model is that it has not been constructed on a scientific basis, but that it has been organized around one author's impressions of a collection of anecdotes obtained from dying patients. Consequently, this lack of extensive empirical data makes that the validity of the model depends on a single author's interpretations (Corr, 1993). This criticism is supported by Kastenbaum and Costa jr. (1977) who claim that the Kübler-Ross' theory is "a very narrow and highly subjective interpretation in which observations and intuitions have been expanded into unwarranted generalizations." Further criticism is that there is no concrete evidence to think that there are "in fact only five ways in which human beings cope with anything as fundamental as dying." (Corr, 1993) Both Kastenbaum and Costa, and Corr argue that even if the five different stages of dying are assumed to be true, the Kübler-Ross model presents no evidence that the stages are somehow interlinked and that a patient actually moves from one stage to the next. Ergo, as it can be the case that a person does not need to experience all five different stages in the grieving process, one might wonder if the given five stages are sufficient to describe all the thoughts and feelings that person is experiencing or that additional stages might be determined. The problem is that the stages seem to lack clear definitions. Kastenbaum and Costa mention that "the significance of pre-terminal personality, developmental level, ethnic orientation and other life history factors is not considered, nor are such critical situational factors as the actual disease process, nature of the treatment, and the sociophysical environ-

ment in which the terminally ill person finds himself." (Kastenbaum and Costa jr., 1977).

In response to the criticism of lacking empirical data to support the Kübler-Ross model, studies have been performed to examine if the Kübler-Ross model and other stage theories of grief holds under the scrutiny of empirical research. One study in particular strengthens stage theories of grief through its findings. Known as the Yale Bereavement Study, Maciejewski et al. (2007) performed a longitudinal study with the aim to examine if the patterns of changes of five grief indicators were consistent with the stage theory of grief. The indicators used in the study were obtained from the hypothesized stage theory of grief by (Jacobs, 1993) that is strongly based on the Kübler-Ross model. It should be mentioned that the participants in this study were bereaved individuals that lost a family member or loved one through natural causes instead of being terminally ill patients themselves. The measurement of grief indicators started one month after the loss of a loved one. The results of the study show that the patterns of grief processing correspond strongly with those proposed in the stage theory of grief, both in the description of the stages and in the temporal sequence in which the stages occur. It is interesting that, in terms of absolute frequency, the initial, dominant grief indicator found was not disbelief (or, denial) but acceptance. This seeming diversion from the stage theory might be caused by the fact that awareness of the approaching demise of the patient promotes acceptance of the death after the loss. Also, as the processing of grief starts at the moment of diagnosis and not at the moment of loss, it is possible that the participant has already experienced the other stages. This view is strengthened by the lower frequencies of acceptance in participants who lost a family member whose diagnosis of the terminal illness was within six months before death compared to patients that were diagnosed six months or longer prior to death. In addition, participants who reported receiving the diagnosis within six months before the death displayed a significantly higher frequency of disbelief grief indicators after the loss. However, after rescaling to take into account the progression of each type of psychological response, the order of the grief indicators exactly matched those presented in the stage theory of grief. This result shows through empirical evidence that the grieving process experienced by people matches that presented in the stage theory of grief, giving credence to the model of Kübler-Ross.

3.3 Conclusions

In this chapter, we have presented what exactly constitutes bad news and how such news is expressed in conversations in a medical setting. In order to find a clear and concise definition, we have looked at existing literature and found that defining bad news is difficult due to two factors. The first one is the situation in which the news is presented and the second one is how the news is appraised by the recipient. Also, we have given an overview of the most common difficulties that naturally occur when holding a bad news conversation, for both the sender as well as the receiver. Although these difficulties often affect either the behavior of the doctor (i.e. the sender) or the behavior of the patient (i.e. the receiver) directly, their occurrence

also influences the overall course of the conversation for both interlocutors. As we have shown, most of the difficulties are directly or indirectly related to the internal state of the patient. The content of the bad news often strongly affects the intensity of the patient's emotions, which causes the selection of his conversational behavior to be determined by his emotions rather than by his rational thoughts. Also, not understanding the information provided by the sender affects the thoughts of the patient. Furthermore, when the patient employs a coping mechanism of not wanting to understand the bad news i.e. denial, the progress of the conversation might be severely hindered.

For the doctor, the difficulties related to the patient's internal state are the following:

- It is unclear to the doctor how the bad news will affect the internal state of the patient.
- The doctor has made an incorrect assessment of the internal state of the patient when preparing for the bad news conversation.

Additional situations that can cause difficulties from the perspective of the doctor are when:

- the doctor is uncertain of the prognosis.
- the doctor experiences feelings of failure to face the patient's need.

In order to deal with these difficulties various studies, protocols and techniques have been developed over the years. In section [3.2.2] we have given an overview of some of the techniques that discuss dealing with the difficulties from the doctor's perspective and we have discussed one of the most prominent protocols in detail. Section [3.2.3] provided an overview of models and theories that describe the manner in which patients might deal with receiving bad news. Again we have taken one of the most prominent models and discussed it in detail.

From the information that has been presented in this chapter the following conclusion can be drawn. In order to properly model the conversational behaviors that are performed in a bad news conversation and the selection processes that precede them, the features of the internal states of both interlocutors and all the processes that are associated with them need to be included in the dialogue model. However, in order to gain a more complete understanding of the features of the internal state, the processing that is involved, how the features and the processes are related to each other and subsequently how they are related to the conversational behaviors, we need to study them in much more detail. The next chapter discusses models of the features and processes of the rational part of the internal state (e.g. thoughts and goals), and also covers models that describe how the affective aspects of the internal state can be represented and explains how they operate with respect to conversational behaviors.

4

Chapter 4: Internal state features

A major part in holding a natural conversation is the ability of the listener to provide an adequate and appropriate response to the conversational behavior that the speaker directs at him. The selection of such a communicative response depends on a large set of factors. In a most basic and simple form of conversation, the selection of appropriate responses can quite easily be handled through clear-cut rules or standards that hold for social conversational interactions. Rules that determine that, for example, greetings should be answered by return greetings, questions should be followed by answers and acknowledgments should be given when information is provided by the speaker. However, relying only on such clear-cut rules does not suffice when attempting to realistically model the cognitive processes that underlie natural conversational behavior. Neither does it suffice when the listener is involved in a more complex conversation. In both situations straightforward rules continue to have some influence on the selection of response behavior, but they are not the driving factor behind the behavior selection process. Cognitive factors such as emotions, social features or listener's goals and intentions will play a more predominant role in the selection of appropriate behavior, making that process much more complex. This holds especially when the focus of the conversation concerns the listener's personal interests such as in a bad news conversation.

In this chapter we present an overview of models and theories that discuss how different parts of a human's internal state can be represented. In the rest of this thesis we will use the terms *cognitive feature* or *internal state feature* to indicate such an individual part. We also use the term *internal state element* to indicate representations of internal state features in our cognitive dialogue model. In section [4.1] we will discuss the BDI model that can be used to represent the core features of an internal state. Sections [4.2] and [4.3] focus on theories about the representation and workings of emotions, which describe how emotions are formed and how emotions influence the selection of appropriate response behavior. These processes are known as *Appraisal* and *Coping* respectively. In section [4.4] we discuss several social features that play a role in holding bad news conversations and which are included in the dialogue model. In section [4.5] we present the conclusions we have drawn from this chapter.

4.1 Beliefs, Desires & Intentions

The dialogue models that have been discussed in chapter [2] can be used as a basis upon which different types of dialogue systems can be constructed. Because the bad news conversations that are focused on in this thesis are quite complex and because the speaker needs to take into account both his own internal state and the internal state of his interlocutor we have decided to use an agent-based approach to dialogue modeling.

A large number of agent-based dialogue systems that have been constructed over the years use the Belief-Desire-Intention **(BDI)** model (Bratman, 1987) to represent the internal state of the agent. Optionally, the agent might use the BDI model to form a user-model of the human interlocutor. The BDI representation of the **beliefs, desires** and **intentions** of the agent (and possibly its human interlocutor) is based on Bratman's folk-psychological theory on human practical reasoning about situations that might be brought about in the future. The BDI model describes what a person or agent knows and wants, expressed in terms that are intuitive for humans to grasp, i.e. **beliefs, desires** and **intentions**.

One focus of the BDI model are the **desires** of the agent. The **desires** are expressions of the motivational state of the agent and indicate which state the agent wants the environment to be in. Examples of these are "I want to be home" or "I want to get better". In the context of conversations, desires often include the agent being in a particular state of mind (e.g. to know or belief something) or about changing the internal state of the agent's interlocutor, e.g. "I (the agent) want to let you (the interlocutor) know I'm not feeling very well." Quite often desires are taken to be identical to an agent's goals in order to make them more tangible. Yet there are subtle differences between them. **Desires** are the representation of an agent's will, while **goals** are used to represent the desired state of the environment. As such the two terms are indissolubly connected to each other as goals follow from desires. However, there are other factors that function as source of an agent's motivation which can lead to goals, namely obligations and norms (Dignum et al., 2002). Furthermore desires differ from goals in the fact that **desires** can be contradicting, while the **goals** of an agent need to be consistent with each other. For example, in addition to having the desire "I want to be home" the agent can also have the desire "I want to be at the beach". Obviously these two states can not be fulfilled at the same time, but they can both be desired by the agent. Our cognitive dialogue model represents the internal state and the processes of a patient involved in a bad news conversation. In the model, we have only included the goals of the modeled patient and not the source of his motivations, as the goals are sufficient to study how and why specific conversational behaviors are selected over others. The process of desire-formation lies outside the scope of this study.

While an agent can have several goals at the same time, he can only pursue one of them at a given moment. The agent needs to commit to bringing about a single goal in order to avoid performing behaviors that have contradicting outcomes. Although the goals might not be contradicting, the actions performed to achieve them might be. Committing to achieve a particular goal turns that goal into an **intention**. For

example, a patient (i.e. the agent) wants (i.e. has the goal) to get information about his state of health from a doctor, but also wants to inform the doctor about his allergy to some medication. While these goals do not interfere with each other, aiming to fulfill them cannot be done at the same time, because this would lead to a very incoherent conversation. Instead the agent chooses one of the two goals to fulfill first and commits to achieve this goal. **Intentions** should be kept until the goal is obtained or until the goal becomes unobtainable.

For each goal in the internal state, the agent system makes a plan to see it fulfilled. In the dialogue model these plans consist of sequences of conversational behaviors that aim to bring about the desired states, i.e. the goals. The process of selecting appropriate behaviors consists of the system evaluating and combining cognitive features that are present in its internal state. According to the BDI model, the agent commits to its intentions based on the knowledge it possesses. This knowledge is represented in the internal state as **beliefs**, i.e. statements about what the agent holds to be true. This includes statements about the environment, but also about the internal state of the interlocutor (i.e. Theory of Mind). The formation of a particular plan (i.e. a sequence of conversational behaviors) depends on whether certain beliefs that are associated with those behaviors hold in the internal state. What these specific beliefs are depends on the conversational behaviors. The selection of a conversational behavior cannot occur when the beliefs that are associated with that behavior do not hold. However, selecting appropriate conversational behaviors is not only done based on beliefs, but **emotions** and **social features** also influence the selection process.

One of the benefits of the BDI model is that over the years various formalizations have been constructed that allow for reasoning about and validation of the performances of the agent whose internal state is built according to the model. These formalizations range from possible-world formalisms and modal logic to temporal logic and have allowed for axiomatization of some BDI notions such as “belief-goal compatibility” and “goal-intention compatibility”. Examples of such formalizations can be found in (Cohen and Levesque (1990), Rao and Georgeff (1991) and Wooldridge (2000)). Another benefit of using the BDI model to represent the internal state of an agent is that it facilitates a separation between planning behaviors and executing them. This allows for both processes to be represented in different modules in the dialogue model which are continuously active in parallel. Furthermore, as **intentions** are, partially, based on existing **beliefs** the agent has, it is quite simple to allow a dialogue model to form intentions (and subsequently to plan the performance of conversational behaviors) that are based on beliefs about the internal state features of its interlocutor.

However, because the agent’s perceptions do not only influence its current beliefs, but also affect its desires and intentions directly, it is more difficult to regulate the processes of selection and performance of appropriate conversational behaviors. In addition, agents constructed according to the BDI model function rationally and thus do not let features such as emotions or social factors influence the selection and performance of their conversational behaviors. The last couple of decades have seen a large increase in research extending the BDI model with emotion models and social features (see for example Ochs et al. (2005), Adam et al. (2006), Pereira et al. (2008)

and Steunebrink (2010)). Formalizing existing emotion models by using BDI concepts might lead to a stronger relationship between psychological theories of emotion and rational cognitive models of internal state representation.

4.2 Appraisal

Natural, complex conversations often consist of more than rational and functional conversational behaviors. Many conversational behaviors also express the **emotions** and various social dispositions, such as **liking** and **rapport** between the interlocutors, or cause them to arise. A conversational behavior of the interlocutor might influence not only the rational components of the agent's internal state (i.e. his **beliefs, desires/goals** and **intentions**) but also his **emotion** and **social states**. Changes to the emotion or social components in the internal state can, amongst others, be caused by evaluating one or more of the following things:

- The content of the interlocutor's conversational behavior,
- The perception of `displays of emotion` or `displays of social states` through facial expressions, prosody, gestures or body posture.
- The manner in which turn-taking takes place in the interaction, (e.g. interruptions).
- Knowledge (i.e. **beliefs**) about a particular state of the environment.

Note that the state of the environment mentioned in the last bullet point also includes the internal states of both the speaker and his interlocutor. For example, if the speaker really wants his interlocutor to know something, then performing the act of telling him (and thereby changing the internal state of the interlocutor) might cause the speaker to experience the emotion **Joy**, without the interlocutor having to perform any behavior.

How exactly a person's internal state is influenced is determined by various cognitive processes. The general terms under which such evaluating processes can be grouped are called *cognitive appraisal processes*. In the field of dialogue modeling there is a preference to use one of two specific types of appraisal processes. The first type of process, *linguistic appraisal*, deals with the evaluation of performed conversational behaviors with respect to the units, structure, and form of the linguistic expressions in the conversational behaviors. Its foci include which phase the dialogue is in (e.g. start, middle or end), which of the interlocutors currently has the turn and what type of dialogue act is being performed. This information is used to manage the dialogue and to determine what the most appropriate type of response behavior is. The second type of appraisal processes that is widely included in dialogue models is that of *affect appraisal* or *emotion appraisal*. These processes evaluate how the different features expressed in conversational behaviors influence the emotion state of a person. In the following sections we discuss various affect appraisal theories and subsequent emotion models.

4.2.1 Lazarus & Folkman

At the beginning of appraisal research in the sixties, Lazarus and Folkman (1984) developed their theory of cognitive appraisal. The theory of cognitive appraisal describes the relationship between a person and the social and physical environment they are in. This is referred to as the person-environment relationship. Included in the environment are also the people with which a person might have a conversation. Lazarus and Folkman divide cognitive appraisal into two separate processes: *primary appraisal* and *secondary appraisal*. People use *primary appraisal* to perceive and interpret the situation of the environment with respect to their personal interests or goals. *Secondary appraisal* is the process of evaluating the potential responses that can be performed as a reaction to the situation. In addition to the primary and secondary appraisal, Lazarus and Folkman's theory contains a second component that deals with the *process of coping*. This *coping process* covers the mechanisms of actively dealing with the situation in order to repair or maintain the state of the person-environment relationships. More information on coping mechanisms is provided in section [4.3].

Primary appraisal distinguishes three types of outcomes into which a situation can be categorized: (1) *irrelevant*, (2) *benign-positive* and (3) *stressful*. When the situation has no implication on the well-being of the person, it falls into the first category; *irrelevant*. The second category contains those situations that are perceived as preserving or enhancing the well-being of the person or that promise to do so. The third category contains those situations that are perceived as to negatively influence or threaten the state of the person. Interpretations of the situations in this third category are called *stress appraisals*. As stress appraisals are strongly related to the situations described in a bad news conversation, we describe them in more detail.

Stress appraisals can be differentiated into situations that describe *harm/loss, threats* or *challenges*. Harm/loss situations indicate that some form of damage has already been sustained by the person. This might include a severe injury or illness, the realization that the person's self- or social esteem has been damaged or damage to or loss of a loved one. As the events that brought about the situation already have occurred in the past there is no way to undo these events through coping responses. Therefore, coping responses to harm/loss will always focus on the current situation. Particularly applicable in these situations are the five stages of dealing with bad news as described in section [3.2.3].

Threat situations are those where the damage to the person has not yet occurred but is anticipated. As such, threat situations are connected with harm/loss situations, as harm/loss situations will also have negative implications for future situations. An important feature that distinguishes threat situations from harm/loss situations is that in the case of threat situations anticipatory coping can be used, whereas this does not hold for harm/loss situations. Anticipatory coping allows people to deal with expected threats to their well-being in advance, something that cannot be done if the events that brought about the situation already have occurred. For example, after having been told that their health will deteriorate in the future, people may take steps to avoid that situation. In the context of a bad news conversation this may include asking advice from the doctor about possible procedures. Such procedures depend on the nature of the hazardous situation, but examples include surgical removal of

tumors, chemotherapy, use of medication or the transplantation of an organ.

Challenge situations have much in common with threat situations in that they too can use anticipatory coping strategies to deal with the stressful situations. The main difference is that challenge appraisals focus on the potential for gain or growth of the person inherent to the situation. Contrary, threat appraisals center on potential harms the person might experience. As a result challenges are characterized by pleasurable emotions such as eagerness, excitement and exhilaration whereas threat (...) is characterized by negative emotions such as fear, anxiety and anger. (Lazarus and Folkman (1984), p.33). It is important to note that a single conversational behavior can be appraised multiple times on how it affects the person's well-being, for example both as a challenge and a threat. Subsequently, evaluation of a conversational behavior might yield more than one emotion to be stimulated. However, Lazarus and Folkman seem to focus more on the processes of appraisal and coping than on the specification of different emotions.

Whenever primary appraisal classifies a situation as being stressful and it falls into the categories of either threat or challenge or both, people are inclined to actively resolve the situation. Subsequently, secondary appraisal processing is performed, which consist of assessing the possibilities of how the situation can be dealt with. Note that secondary appraisal is more than simply making a list of possible response behaviors. It is "a complex evaluative process that takes into account which coping options are available, the likelihood that a given coping option will accomplish what it is supposed to, and the likelihood that one can apply a particular strategy or set of strategies effectively." (Lazarus and Folkman (1984), p.35). In order to form these coping options, the secondary appraisal process takes into account a person's current beliefs and intention, but also the newly stimulated emotions.

4.2.2 Scherer

An alternative approach has been proposed by Scherer (1984), who has developed the "component process model of emotion" to represent the processes of forming and maintaining emotions. This model consists of multiple components, one of which is the cognitive appraisal component. The other components include physiological activation, motor expression, motivational tendencies, and subjective feeling state. The cognitive appraisal process in Scherer's model evaluates the conversational behavior with respect to several aspects: its novelty and unexpectedness, its intrinsic pleasantness, its congruence with goals, the coping possibilities, and its compatibility with norms. Furthermore, similar to the theory of Lazarus and Folkman (1984), Scherer also sees the importance of a coping process to produce appropriate response behaviors.

Scherer's appraisal consist of three separate processes that are closely interlinked with each other (Leventhal and Scherer, 1987). Two of the three processes are automatic and can be considered to be part of a third conscious process. The first automatic process appraises behavior at the sensorimotor level. This process focuses on automatic, unconscious evaluation of the stimulus behaviors and induces reflex responses, such as pain leading to sadness. The second automatic process performs appraisal at the schematic level. This includes the automatic mapping of the per-

ceived behavior to behavior patterns previously learned, which are already present in the internal state. Appraisal at the schematic level induces coordinated or learned responses. At the conceptual level, a third appraisal process known as conscious appraisal, becomes active when the two underlying automatic processes cannot determine an appropriate emotion response to the stimulus behavior. The conceptual appraisal process uses knowledge stored in the internal state (i.e. **beliefs**) to evaluate the stimulus behavior consciously. When the evaluation of the stimulus behavior warrants an emotional response (as determined by the aspects mentioned in the first paragraph) the appropriate **emotions** are generated. If the same type of stimulus behavior occurs over and over, this response will become automatic.

With respect to the specification of different types of emotions, Scherer suggests that there may be as many different emotions as there are appraisal outcomes. Which emotions are stimulated based on which appraisals does not become very clear in Scherer's work on appraisal.

4.2.3 Ortony, Clore & Collins

The psychological model of emotion by Ortony et al. (1988) (commonly known as the OCC-model) has been used frequently in the past, either as starting point or inspiration. Compared to the two models presented in the previous sections, the OCC-model focuses more on how the appraisal of behaviors relates to emotions and less on how the appraisal processes function. The model provides a theoretical, structural approach to emotions, containing detailed observations about emotions and their organization. Instead of specifying individual emotions, the OCC-model describes emotion types. The idea is that individual emotions are best represented as being part of substantially independent groups, based on the nature of their cognitive origin. Each emotion type can be thought of as representing a family of closely related emotions. Emotions of the same emotion type might for example differ in terms of their intensity, but still be quite alike. An example of this is "contentment" compared to "euphoria", both of which are part of the emotion type **Joy**. Part of the description of the emotion types are specifications of the conditions that elicit the emotions they express and any variables that may affect the intensities of these emotions. The OCC-model specifies 22 different emotion types as possible reactions to three kinds of stimuli that can be appraised. These 22 emotions types are shown in figure [4.1].

According to the OCC-model, appraisal can be done on the consequences of events, the actions of agents and the aspects of objects. A complete overview of the emotion types of the OCC-model is included in appendix [B].

When used in a dialogue system, the process of appraisal looks for eliciting conditions and variables in the observed conversational behavior that might be associated with a particular emotion type. If found, the conditions and variables are evaluated and subsequently an emotion is elicited. The eliciting conditions are determined based on the disposition a person has with respect to the stimulus. These dispositions can be seen as valenced reactions with respect to the stimulus. They can vary from *pleased / displeased* in cases of CONSEQUENCES OF EVENTS, *approving /disapproving* in cases of ACTIONS OF AGENTS and *liking / disliking* with respect to ASPECTS OF OBJECTS. In addition to the dispositions, elicitation depends on the value of the variable

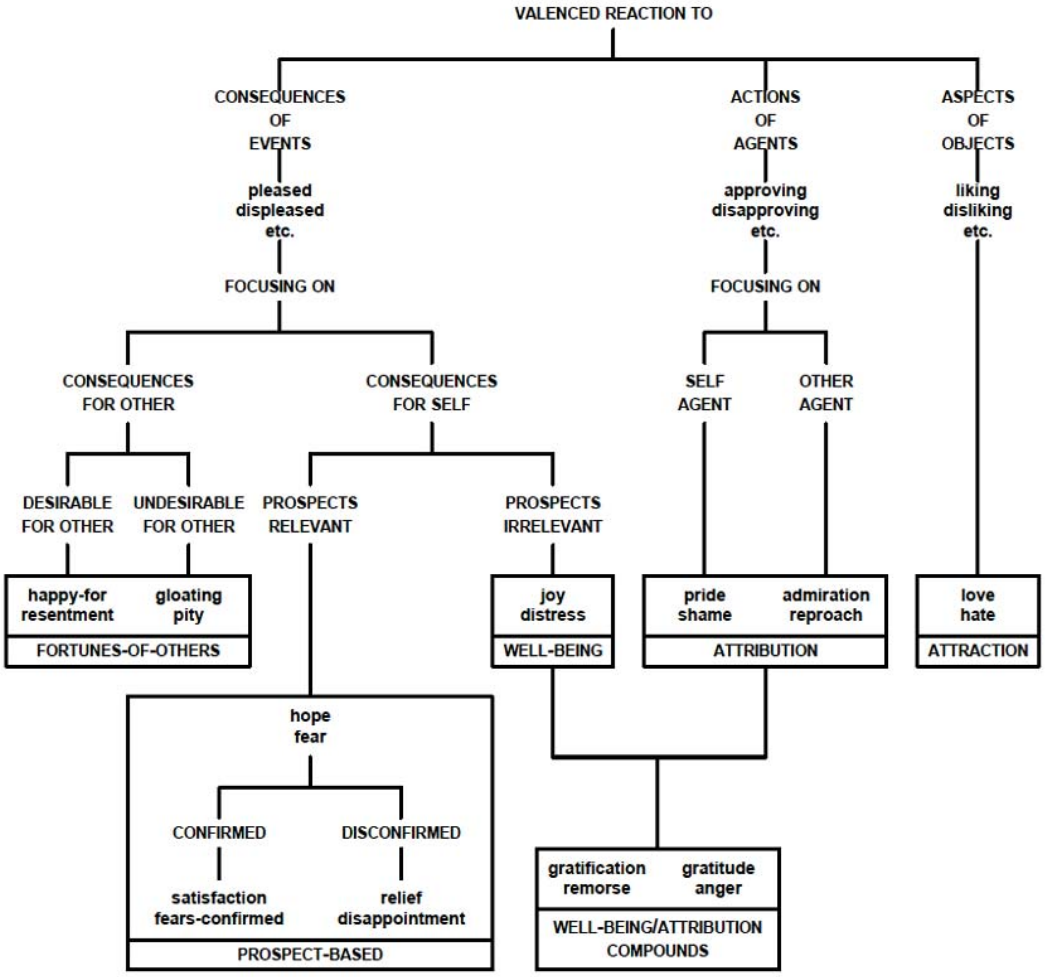

Figure 4.1: A view of the emotions of the OCC model. (Ortony et al. (1988), p.19)

associated with the emotion type. Each emotion type has a principal variable that is local to the group of emotions that make up the emotion type, affecting their intensities and allowing them to be elicited. Furthermore, the OCC-model also specifies global variables that can affect the intensity of all emotions, regardless of which emotion type they are part of. For future references, we call the variables, both global and local, that influence the intensities of the emotions ***'appraisal variables'***.

When appraising the CONSEQUENCES OF EVENTS, the condition that might elicit an emotion depends on the ***desirability*** of the event with respect to the person's goals, in addition to his disposition. The value of the local variable ***desirability*** can either be *desirable* or *undesirable*. For example, if a person has the goal to walk in the park, the event "the sun starts shining" will be appraised as *desirable* (it is positive with respect to the person's goal). Combined with the disposition of being pleased, the emotion type that is consequently elicited is **Joy**. The eliciting condition associated with this emotion type is "being pleased about a desirable consequence of an event."

In the same way, the appraisal of the actions of agents takes into account the local

variable ***praiseworthiness*** with respect to a person's social standards. (More details on social standards are presented in section [4.4]). If the person's standards dictate an action is *praiseworthy*, for example getting a good grade on a test, performing that action elicits an emotion from the **Pride** emotion type. The disposition associated with the eliciting condition in this case is APPROVING OF THE ACTION OF THE AGENT, i.e. approving of one's own praiseworthy action.

If aspects of objects are being appraised, elicitation of emotions may be achieved based on the local variable ***appealingness***. The ***appealingness*** of an object is measured with regards to a person's attitudes towards or preferences for ASPECTS OF THAT PARTICULAR OBJECT.

Appraisal of stimuli of the three major categories are thus primarily based on ***desirability, praiseworthiness*** and ***appealingness***. Note that the dispositions needed for the elicitation of emotions remain the same for each of the three types of stimuli. In addition to these local variables, the OCC-model contains other local variables. These other variables that lead to elicitation of emotions are for example the ***likelihood of the consequences of an event coming to pass*** (leading to **Fear** or **Hope** emotion types), whether the ***consequences of the event apply to others*** and then what the ***desirability for others*** might be (for emotion types such as **Happy-for** or **Pity**), or the ***confirmation of the consequence of an event*** (for emotion types such as **Relief** and **Disappointment**). In addition it can sometimes occur that a combination of different variables elicits an emotion. For example, the emotion type **Anger** arises when a person evaluates an action of an agent as *blameworthy* and the consequence of the (accompanying) event as *undesirable*. The combination of related stimuli might also elicit multiple emotions at the same time. One might be happy the sun is shining, but at the same time fearful that one might get a sunburn.

An example of a global variable is the ***unexpectedness*** of the stimuli, which can have the value of either *expected* or *unexpected*. This variable can influence the intensity of emotions elicited by CONSEQUENCES OF EVENTS (e.g. finding a coin of the street), ACTIONS OF AGENTS (e.g. a surprise visit from a favorite uncle) or ASPECTS OF OBJECTS (the softness of a piece of wood). Another example of a global variable is ***arousal*** or the ***degree of excitement*** the person experienced before the stimulus was presented.

One must take care to categorize the situation described in a conversational behavior correctly, as it is quite easy to make mistakes. For example, an ACTION OF AN AGENT (singing a song) might receive the utterance "I love it when you sing" as a response. In this case, the ACTION OF THE AGENT was not the focus of the appraisal, but rather the OBJECT (i.e. the song) the action produced.

As mentioned earlier, the OCC-model has been frequently used in the past including in the construction of dialogue models containing emotions. This is because the OCC-model has been formulated in a very structured manner with a limited number of concepts, making it well suited to be integrated into another formalized model. The classification of the emotion types is very clear and detailed, and covers a wide range of different emotions. Also the OCC-model specifies quite clearly the conditions that elicit the emotion. Furthermore, the OCC-model uses concepts that are also prevalent in the BDI model, such as beliefs and goals. The benefit of using these concepts is

that they are well-studied, are already formalized in various studies and have strong explicative power of the person's conversational behaviors and internal state.

However, appraisal of observed stimuli is only one part in the process of dealing with emotions. When elicited, emotions also often lead to the performance of an appropriate response behavior. The question which response behavior should be performed and how it should be performed is something that is covered by the coping process. In the next section, we present explanations of the coping process.

4.3 Coping

Coping is a term that has become quite well known and well used with the rise of stress related research in the last few decades, both in the academic world and in everyday life. It is the process of performing those behaviors that best deal with the threatening or challenging situation at hand, or as Lazarus and Folkman define it: "the constantly changing cognitive and behavioral efforts to manage specific external and / or internal demands that are appraised as taxing or exceeding the resources of the person." (Lazarus and Folkman (1984), p.141). As such, the coping process has a large influence on the selection of behaviors in stressful situations and it seems prudent to examine how the coping process operates. Furthermore, examining the coping process provides information about which type of features of the internal state are used in selecting appropriate response behaviors.

This leads to a range of questions about coping, such as: 'What are the internal state features that play a role in the coping process and how do they do that?', 'To what extent are people consistent in their coping behavior in a large diversity of stressful situations?' and 'How can we discriminate between different coping behaviors, such that we can categorize them?' In this section we aim to find answers to these questions, as they can assist us during the constructing our dialogue model, particularly the component concerned with handling emotions and conversational behaviors associated with them. In addition, we argue that the process of coping plays a large role in selecting appropriate conversational behaviors.

As mentioned in section [3.2.3], the mental approaches to bad news taken by recipients vary based on the contents of their internal states. In the BDI model, a patient's thoughts about the bad news situation are represented by his beliefs, and his desires and goals represent how he wants the situation to be. Together with his emotions, the beliefs and desires form the patient's disposition towards the situation and determine his intentions and subsequently his conversational behaviors. The manner in which the patient copes with the bad news situation (i.e. his coping strategy) is expressed by the combination of specific conversational behaviors.

In order to determine whether the conversational behaviors performed by the patient adhere to his coping strategy or is just an unstructured sequence of behaviors, it is helpful to know if people are consistent with themselves in the way they deal with various, stressful situations. In addition, it helps to know if there are consistencies in the conversational behaviors different people perform in similar bad news situations.

Consistency of coping behaviors depends on the person as well as the situation that is appraised. According to Folkman and Lazarus (1980), the consistency with

which people select and perform their coping behavior might be affected by ***person variables*** and ***situation variables***. ***Person variables*** are variables that are determined by the current internal state of the person coping, while ***situation variables*** are determined by the situation that is appraised. If coping is determined only by person variables, then similar variable values in the internal state will result in the same coping behavior, regardless of the situation the person is in. If the coping behavior is primarily determined by situational variables, then the coping behavior performed in different situations should be different even if the person variables are the same and thus will not be consistent when the person is experiencing the same feelings. This raises the question of how the person variables and the situation variables are related to each other. It seems that the state of the situational variables plays a role in the construction of the person variables (i.e. the internal state of the person) about the situation.

If conversational coping behaviors performed by different people in similar situations are consistent, or if different persons are consistent with themselves in situations that are alike, then we can group these conversational behaviors and label them as a particular coping strategy. Grouping conversational behaviors and their associated internal state features helps when making assumptions about the internal state features of a speaker when he performs a new set of conversational behaviors that is similar to the one already grouped.

The question is how conversational coping behaviors can be grouped together and into which categories. Luckily, efforts to detect and categorize coping behaviors that are alike have been made in the past. This has led to a range of approaches of measuring coping behaviors.

4.3.1 Measuring coping

Approaches to detection and subsequently categorization of coping behavior have historically been based on three broad perspectives (Folkman and Lazarus, 1980). The first of these perspectives conceptualizes *coping as defensive or ego processes*, where the defensive or ego responses of a person are often assigned through evaluation of the person's overall functioning or his adaptational outcomes (i.e. his behavioral responses). The second perspective has conceptualized *coping as a personality trait*. Traits differ from defensive processes in that "they refer not to a defensive response per se, but to the dispositional or personality attribute that leads to the response." (Folkman and Lazarus (1980), p.221). This entails that coping is attributed to a person based on interpretation of his internal state, rather than through interpretation of his behaviors. The third perspective catches the concept of *coping in terms of the special demands of specific kinds of situations* (such as illness, natural disasters and bereavement) and is used by researchers who are situation-oriented.

According to several studies that use this situation-oriented perspective (e.g. Weisman (1976), Bulman and Wortman (1977) and Folkman and Lazarus (1980)), coping strategies can be grouped into functional categories such as: strategies for maintaining or restoring interpersonal relationships, seeking information, feeling better, maintaining self-esteem, making good decisions and reduction of anxiety. However, these studies are usually designed to detect the coping strategies in specific and unusual

situations, such as serious illness or severe injury. Consequently, most of the coping strategies described in each particular situation cannot be generalized to other contexts or situations. However, some of the coping strategies (e.g. **engaging in selective perception** and **making positive comparisons**) were found in several different situations (Pearlin and Schooler, 1978), suggesting that there are coping strategies that are generally applicable. This means that when coping behaviors are observed across several different situations, both consistency and variability in coping strategies can be determined.

Each of the three perspectives mentioned poses difficulties in dealing with determining coping strategies for various reasons. For example, **detection of defensive processes** is based on information about how the person functions, i.e. based on the behaviors he performs. Although these behaviors are closely related to the coping process, these two things are not the same. Because coping strategies are determined by evaluating the observed behaviors, the found coping strategies cannot be used again to explain the subsequently performed behaviors, as this would be circular reasoning. Furthermore, it is difficult to get an adequate interrater reliability for classifying the defensive processes (Folkman and Lazarus, 1980). The problem with determining coping strategies by using traits is that decisions are made based on the assumption that the coping behaviors of people are consistent across situations. If traits are used to determine the consistency of coping behaviors, people hardly ever seem to be consistent and when they are, it is only in certain very specific situations. As a consequence, traits are poor predictors of coping processes (Cohen and Lazarus, 1973). The problem with determining coping strategies based on situation-oriented coping processes is that these processes are also too situation specific.

4.3.2 Ways of Coping

In an attempt to circumvent these problems, Lazarus and Folkman take another approach to conceptualizing coping (see Lazarus and Folkman (1984), Folkman and Lazarus (1980) and Folkman and Lazarus (1985)). They propose to conceptualize coping as conscious efforts to manage stressful demands, regardless of the outcome of the response behavior. This approach to coping is called *cognitive coping* (Lazarus, 1999). The separation between the adopted coping strategy and the outcome of the associated coping behavior is made in order to avoid confounding the two (Lazarus and Folkman, 1984). Folkman and Lazarus perceive cognitive coping as one of two components that mediate the ongoing relationship of reciprocal actions between the person and the environment, the other component being *cognitive appraisal*. Cognitive appraisal has already been discussed in section [4.2.1].

Both processes continuously influence each other throughout interactions between two interlocutors. Consequently, the answer to the question whether coping determines appraisal or appraisal determines coping is "provisional depending upon where one interrupts the ongoing dynamic relationship." (Folkman and Lazarus, 1980). For example, a situation that is appraised as stressful will be dealt with through the performance of coping behavior. This will lead to a new situation that will subsequently be appraised again. As the appraisal process might be considered to be a transformation of ***situational variables*** into internal state variables (i.e. ***person variables***)

and the selection of appropriate response behaviors the reverse of this, handling the situation (i.e. coping) is thus done with respect to both situation and person variables.

Based on this approach to coping and appraisal, Lazarus and Folkman have developed the 'Ways of Coping' (see Lazarus and Folkman (1984), Folkman and Lazarus (1980) and Folkman and Lazarus (1985)); one of the most established manners to measure and categorize coping behaviors. The Ways of Coping method consists of a checklist of measurement items describing a wide range of behavioral and cognitive coping strategies that a person might use in a specific stressful situation. After revising, the checklist consists of 66 self-report measurements with answers distributed over a 4-point Likert scale. The offered coping strategies have been derived from various sources in previous coping literature, including work about defensive coping, information-seeking, problem solving, palliation, inhibition of action, direct action, and magical thinking. Based on data acquired using the checklist, Folkman and Lazarus describe two characteristics that can be assigned to types of coping such that they can be categorized (see Folkman and Lazarus (1980) and Folkman and Lazarus (1985)).

On the one hand, there is *problem-focused coping*; strategies of behavior that focus on resolving the problem or doing something that alters the sources of the stressful situation. Problem-focused coping can be further divided into two major groups of strategies: those *directed at the environment* and those *directed at the self*. The first group includes strategies such as **altering environmental pressures, barriers, resources, procedures etc**. The second group includes strategies that involve altering the coping person's motives and cognitive abilities, such as **shifting the level of aspiration, finding alternative channels of gratification, developing new standards of behavior**, or **learning new skills and procedures**. While strategies in the second group are not typical problem solving techniques, they are used to help the person to manage or deal with the problem.

On the other hand, there is *emotion-focused coping*. Emotion focused coping is aimed at the regulation of (possible negative) emotions that are associated or caused by the stressful situation, in an effort to maintain emotional equilibrium. This includes cognitive strategies such as **avoidance, minimization, selective attention** and **positive comparison**. Emotion-focused coping primarily deals with internal emotional states that follow as a result of the stressful situation and not directly with the situation itself.

Taking the *problem-focused / emotion-focused* distinction into account during the analysis of data acquired from the Ways of Coping checklist, Folkman and Lazarus produced eight scales into which coping behaviors can be divided (Folkman and Lazarus, 1985). These scales include one problem-focused scale, six emotion-focused scales and one scale that indicates both problem- and emotion-focused items. The eight scales are:

- **Problem-focused coping** (e.g. adjusting the environment or one's disposition)

Emotion-focused coping:

- **Wishful thinking**
- **Distancing** (e.g. avoidance)

- **Emphasizing the positive** (e.g. positive comparison)
- **Self-blame**
- **Tension-reduction**
- **Self-isolation**

Mixed problem- and emotion-focused coping:

- **Seeking social support**

The skewed division between problem-focused coping and emotion-focused coping might be explained by the fact that the number of problem-focused strategies that are applicable across different situations is relatively small compared to the number of emotion-focused forms of coping (Lazarus and Folkman (1984), p.153). One way of dealing with this is by setting a broad scale of coping strategies into which a large variety of behaviors will fit.

Furthermore, Folkman and Lazarus (1985) indicate that, similar to the possibility that one stressful situation leads to multiple different appraisals, so can it be that the same stressful situation elicits both types of coping strategies. However, in such a case either one or the other strategy is dominant depending on the perception and interpretation of the situation by the person. If the person feels that something constructive can be done about the situation, problem-focused strategies will be prevalent, whereas if the person feels that the stressful situation is something that must be endured, emotion-focused coping strategies will dominate. Results (Folkman and Lazarus, 1980), (Folkman and Lazarus, 1985) show that the percentage of people employing both problem-focused coping and (at least one) form of emotion-focused coping at the same time in different situations is very high compared to the proportion of people that do not.

Folkman and Lazarus (1980) also mention several other factors that might prove useful in the measurement of coping but do not take them into account during their studies. Such factors include, but are not limited to, whether coping is directed at oneself or at another, whether it is active or passive and whether coping produces a solution or merely a deferral of solution.

4.3.3 Carver et al.'s criticism

While the division into problem-focused and emotion-focused coping may be a good starting point in the measurement and categorization of coping strategies, it is not sufficient. Carver et al. (1989) argue that the distinction is too simple. They state that several different coping strategies need to be considered in order to make a good assessment of how people deal with stressful situations. According to Carver et al., researchers generally consider strategies other than problem-focused coping as variations of emotion-focused coping, as Folkman and Lazarus have begun to do (Folkman and Lazarus, 1985). However, these variations can be very different from each other, even to the extent that they are inversely correlated. If all are regarded as being part of the emotion-focused coping category, this might be cause for problems.

For example, some emotion-focused responses to stressful situations involve **denial**, while others involve **positive reinterpretation of events** and still others involve the **seeking of social support**. Obviously, each of these responses is very different

and the success of the outcomes they produce might vary, possibly depending on the specific situation. Instead of just distinguishing between emotion-focused coping and problem-focused coping a more extensive collection of categories needs to be formed. After all, a good distinction between different strategies makes it easier to measure coping behavior.

Furthermore, also problem-focused coping needs to be examined more closely. While seemingly a single behavior, problem-focused coping can possibly be divided into several distinct activities: e.g. **planning, taking direct action, seeking assistance, screening out other activities** or sometimes forcing **oneself to wait before acting**. Each of these response behaviors can be measured individually. The approach Carver et al. (1989) take characterizes this in a broader perspective: in order to make a good assessment of coping, one needs to look at the factors that make up coping and measure those factors individually. This approach is taken in order to circumvent three major problems Carver et al. (1989) regard other measurement approaches to have.

First, while there is a large variety in what different measurements assess, none of the existing measurement approaches cover multiple specific domains.

Secondly, categories determined by the measurements often display a lack of focus. One of the possible causes is a lack of knowledge about the motivation for selecting and performing a particular coping behavior. For example, consider this item from the 'Ways of Coping' scales: "Took a big chance or did something risky." A risky act might be done for any of several diverse reasons. Doing something risky might mean something such as taking drugs or driving recklessly to avoid thinking about the stressor. Alternatively, it might mean taking action that is unlikely to be successful, but that if successful would solve the problem, such as undergoing a dangerous operation. (Carver et al., 1989) Alternatively, the lack of focus may be caused because the description of a category is ambiguous. Take this example from the 'Ways of Coping' scales: "I did something which I didn't think would work, but at least I was doing something." (Carver et al., 1989). It is unclear which fact in this description is more important: the fact that something is being done or the fact that the person did not think the action was going to work.

The third problem with earlier measurement approaches is that a large portion of their categories are derived through empirical research rather than through theoretical methods. More specifically, the items used to measure coping behaviors are often chosen because they are diverse and representative examples of potential coping responses, not because they represent theoretically interesting categories of coping. Identification of the characteristics that compose these categories is often done through factor analysis. This causes the found categories to be somewhat loosely and post hoc related to theoretical principles.

4.3.4 Carver et al.'s approach (COPE)

Carver et al. (1989) propose a measurement method that is developed through a more theory-based path, taking two existing models as guidelines: Folkman and Lazarus' 'Ways of Coping' and an earlier model of Scheier and Carver (1988). Similar to Folkman and Lazarus, Carver et al. (1989) also propose a classification of coping strate-

gies. To this end, they developed the 'COPE' inventory of behaviors, from which they could take measurements of coping. They determined thirteen conceptually distinct scales or strategies. Contrary to earlier measurement methods some of the strategies were not only derived from empirical sources but were also based on theoretical arguments about the functional properties of coping strategies. Other strategies have been included because previous research has indicated that they are sound or that they obstruct adaptive forms of coping. The following strategies have been determined by Carver et al. (1989):

- **Active coping:** This coping strategy deals with the situation by making the person who is coping take active steps to remove or circumvent the stressor or to ameliorate its effects. It consists of initiating direct actions or increasing the effort of actions that are already being performed.
- **Planning:** This strategy involves coming up with actions that deal with the stressful situation. It focuses on which steps must be undertaken and in what order. It is important to note that planning occurs during the secondary appraisal phase instead of during the coping phase as most coping strategies.
- **Suppression of competing activities:** It is also possible that the person deals with the stressful situation better when he directs more attention and focus to it. To that end he may suppress other actions and processes that are interfering with the coping process. This may include actions that impede the coping attempt or information processing channels.
- Another approach to dealing with a stressful situation is to **exercise restraint on actions and processes**. Restraint coping is performed by not taking action, but to wait until a more appropriate time to act upon the situation occurs. This is both an active coping strategy, as the person effectively tries to deal with the situation, as well as a passive coping strategies, in the sense that to restrain from coping implies not to do anything.

These four strategies are more or less similar to what Folkman and Lazarus have termed *problem-focused coping* but are more extensive and specific. Relevant to problem-focused coping are the strategies of **seeking out social support**. Carver et al. have made a distinction between two types of seeking social support in their classification, because, according to Carver et al., seeking out emotional social support inherently has two functions.

- On the one hand there is **seeking social support for instrumental reasons**. This focuses on seeking advice, assistance or information to deal with a stressful situation. This coping strategy is *problem-focused*.
- **Seeking social support for emotional reasons** is obviously a form of *emotion-focused coping*. Seeking social support for emotional reasons focuses on getting moral support, sympathy, or understanding.

While Carver et al. have distinguished these two strategies because they are conceptually different, they have also found that seeking social support for instrumental reasons and seeking social support for emotional reasons often co-occur.

- Related to Seeking social support for emotional reasons, is the strategy of **focusing on and venting of emotions**. This strategy can have a functional goal, such as grieving in a period after the loss of a loved one. However, this coping strategy might not always be beneficial. For example, focusing on one's negative emotions (especially for a long time) can impede adjustment to the situation that caused the emotion in the first place.

According to Carver et al. the following two strategies of coping might also be dysfunctional with respect to dealing with a stressful situation. The first strategy is called **behavioral disengagement**, while the second one consists of behaviors that result in **mental disengagement**.

- The **behavioral disengagement** strategy contains coping behaviors that are aimed at reducing one's effort in dealing with the stressful situation. This coping strategy can even lead to rejecting the selection of certain behaviors and abandoning goals with which the stressor is interfering. The behavioral disengagement strategy of coping is related to the Denial stage of grieving, which is discussed in section [3.2.3].
- Closely related to the behavioral disengagement coping strategy is the **mental disengagement** strategy. Behaviors that fall into this strategy are aimed at distracting one's mind from the stressful situation. Examples include escaping into any mental activity that veers away from the situation, such as daydreaming, cognitive work, sleeping etc. Similar to behavioral disengagement, mental disengagement is related to Denial.

All the coping strategies that have been mentioned so far are based on theoretical approaches and ideas about coping. The remainder of Carver et al.'s coping strategies are based on empirical precedents that suggest that they are also important. Carver et al. (1989) claim that although the remaining strategies are not grounded in theory, links can be drawn from each of these coping strategies to various kinds of theoretical principles.

- One of these empirically based coping strategies is **reinterpretation and growth**. This coping strategy focuses on managing the emotions instead of dealing with the stressor that was the cause for these emotions. This can be done for example by taking a new perspective towards the stressor, thereby possibly altering the appraisal of the situation.
- Furthermore, Carver et al. constructed the strategy **denial** as a coping strategy. Coping behaviors that follow the denial strategy aim to try and deny the presence of the stressor or the situation. While denial is a basic, natural strategy of responding to a situation, it often is counterproductive as it does not do away with the stressing factor. It is even possible that by adhering to the denial strategy, the situation is allowed to escalate making it increasingly harder of coping successfully in the long run. This strategy is present in the stage theory of grieving, see section [3.2.3] for more information.
- The next strategy in the list is the opposite of denial, namely **acceptance**. It is disputable whether the acceptance strategy is a functional strategy. One might

reason that a person who accepts a stressful situation seems too engaged in an attempt to deal with the situation (possibly by employing another strategy of coping). However, one can also reason that the acceptance of a stressful situation as being something that must be endured or adapted to is a passive strategy rather than a functional one. Similar to the denial strategy, the acceptance strategy is present in the stage theory of grieving as one of the stages.

- The last strategy defined by Carver et al. contains behaviors that have **turning to religion** as a strategy. However, as also stated by Carver et al., this strategy can contain a broad variety of different behaviors, each of them related to one of the other strategies, rather than being a concise strategy on it's own. For example, one might turn to religion for **emotional support**, or as a way of **positive reinterpretation and growth** or even to **actively cope** with a stressful situation.

The set of coping strategies described by Carver et al. shows some similarities to the set presented by Folkman and Lazarus, but is more extensive. Furthermore, as mentioned earlier, the coping strategies described by Carver et al. are based on both empirical and theoretical approaches. Although the descriptions of the coping strategies do not provide much information about which elements of the internal state need to hold in order for them to be selected, they do provide us with much insight in what the intents and forms of possible coping behaviors are. This insight is used in the creation of the behavior selection component in the dialogue model and is discussed further in chapter [6].

4.4 Social state

Although appraisal and coping play an important role in the selection and performance of conversational behaviors, they are not the only processes to do so. The features in a person's social state and the processes associated with them also play a significant role in selecting and performing appropriate conversational behaviors.

The social state represents the mental disposition a person has with respect to others on a social level. Represented by the social state are a person's thoughts about the social relations between himself and others, and the impression he has of how his interlocutor perceives him (in a social manner). Examples of elements that can be included in the social state are the **amount of trust** a person has in his interlocutor, **how much he likes** his interlocutor, the **amount of rapport** that exist between them, etc. These elements are influenced by the way the conversational behaviors of the interlocutor are evaluated, for example for the level of politeness of the utterances, and the social roles that are assigned to both interlocutors and the relation that exists between them such as doctor-patient, teacher-student or parent-child.

For the inclusion and representation of the measure of **rapport** in the dialogue model, we look at the work of Gratch et al. (2006a). The work of Gratch et al. is centered around the inclusion of rapport into an embodied conversational agent dialogue system with the aim to improve speaker fluency and engagement. Although the research is well suited for the representation of rapport in the internal state and consequently also for the selection of conversational behavior, the influence on the

form and manner of performance of the behavior is more limited. This is because the influence of rapport on the semantic context of the conversational behaviors is not included in the agent model, but is limited to vocal features of the utterance (e.g. loudness, speed) and non-verbal features (e.g. gaze, gestures and head movements).

For the representation of the **trust** feature of the social state we draw from the research of Bickmore (2003) and Falcone and Castelfranchi (2001). According to the studied research, the social state feature **trust** is most often expressed through what is being said (i.e. the content of a conversational behavior), rather than through any facial displays or prosodic features. In addition, the trust aspect also influences the communicative functions of conversational behaviors and the **intentions** that are associated with them.

4.5 Conclusions

In this chapter we have shown a wide variety of theories and mechanisms that describe ways in which the cognitive state of a participant of a conversation can be represented. Aimed at constructing a cognitive dialogue model that is capable of processing and selecting conversational behavior in a human-like manner, we have focused on theories that have a psychological basis. From the studies presented in chapter [2], we concluded that the most suitable way to represent a cognitive dialogue model is to use agent-based approach that contains an internal mental state. The BDI model presented in section [4.1] offers a solid basis for constructing such a model. We believe the core of the cognitive dialogue model needs to be focused on concepts that represent rational thinking, i.e. possess knowledge about the situation, have goals one wants to see fulfilled and have intentions one pursues. The BDI model allows for all these cognitive features.

Because we are interested in the evaluation and selection processes of conversational behaviors in complex bad news situation, additional features of the internal state, such as emotions and social values, need to be added to the BDI core. The inclusion of such features is likely to make a conversational model more natural and realistic. To that end we have studied various emotion models that focus on the processes of emotion appraisal and coping and have taken a cursory look at theories that describe how social features operate in conversations. The theories and approaches discussed in sections [4.2] and [4.3] have provided a lot of information on how to relate emotions to conversational behaviors and how they affect the pursuit of one's goals. Particularly the OCC model provides a clear view on this process, and subsequently we have decided to use the OCC model in the construction of the emotion state in the dialogue model discussed in chapter [6]. The coping models presented in section [4.3] show how coping strategies influence the selection of conversational behaviors. The coping strategies indicate what kind of behavior is appropriate in particular situations. By selecting the most appropriate coping strategies found in the various models in this chapter and combining them with the coping strategies described in chapter [3], we construct a categorization-base from which conversational behaviors can be selected and which is also described in chapter [6].

The theories described in section [4.4] show that the social relations that exist be-

tween the two interlocutors might influence the features and processes contained in their internal states. In addition, they also indicate that certain properties of conversational behaviors can be related to social state features of both the speaker and the listener. Social aspects are interconnected to almost all other features and processes of the interlocutors' internal states, and although including them adds an additional layer of complexity to the dialogue model, omitting them would devaluate our claim that our cognitive dialogue model represents a human-like way of processing dialogues.

5

Chapter 5: Data collection and analysis

In the previous chapter we presented several methods and theories about how different aspects of the internal, mental state of a speaker in a conversation can be represented and which processing mechanisms are involved in handling them.

In this chapter we examine in detail conversational behaviors that occur in bad news conversations, in order to determine how the internal states of the speaker and the listener, and the processing mechanisms associated with them manifest themselves in conversational behaviors or can be assumed to be at work when the conversational behaviors are produced - either as cause or effect. We present and discuss two analyses that have been performed on the conversational behaviors in a bad news conversation.

We believe that by observing and interpreting the conversational behaviors in a bad news conversation, we can make assumptions about which features of the internal state of the speaker hold at the moment a particular conversational behavior is being performed. Based on the conclusions drawn from the study presented in chapter [4], we have chosen to use the following features to represent the thoughts and feelings that make up a person's internal state: **beliefs** (including expectations about the future), **desires** or **goals, intentions**, an **emotion state** and several interpersonal social relations between the speaker and his interlocutor such as **liking, rapport** and **trust**.

The difficulty of analyzing bad news conversations is that data containing real life medical bad news conversations is very hard to come by due to various practical and ethical reasons. As an alternative, videos of simulated conversations are recorded for use in study and training, using actors to play the roles of coworkers and patients. In health care, such a patient is referred to as a standardized or simulated patient. It is an individual who is trained to act as a real patient in order to simulate a set of symptoms or problems. Simulated patients have been successfully used in medical education, nursing education, evaluation and research. Realistic simulations of bad news conversations are difficult to create because although there are various models and theories how bad news conversations can be performed, these models often do not agree with one another or are insufficient, as shown in chapter [3]. This makes it difficult to set a base-line on how a bad news conversation is normally conducted.

While searching for appropriate and available data that could be used in our study, we came across the Pallium Project [1]. "The Pallium Project is a community of academic, health services delivery, voluntary sector, government and citizen leaders working together throughout Canada to build Hospice Palliative Care (HPC) capacity as part of sustainable health systems and healthy, vibrant communities. Since 2000, the Canadian Pallium Project has promoted collaboration, coordination and communication for access, quality and new capacity in service of the seriously ill and dying at the community-level throughout Canada." (Pallium project description). As part of the Pallium Project, a series of instructional videos about clinical communication in Hospice Palliative Care has been created. These videos are intended to "support a variety of critical reflective conversations in pre-service education and continuing professional development." One of the videos from the series covers the topic of conducting a bad news conversation. A quick assessment of the bad news conversation in the video showed that the conversational behaviors performed in the conversation contained a variety of social and emotional expressions and references to internal state features. The variety and the types of expressions and the inclusion of references to internal state features were suited for our research. Also, the behaviors in the conversation show that the patient enacts some of the coping strategies presented in chapters [3] and [4], and that the doctor is not following a protocol such as **S.P.I.K.E.S.** (see section [3.2.2]). In addition, the fact that the video-clip is widely used as an instructional tool suggests that it is of good quality.

The scenario enacted in the video-clip depicts the following situation. A family doctor is visiting a patient, who has just undergone surgery for colon cancer. Although the primary tumor has been removed during the surgery, the surgeon also found metastases in the patient's intestines. This means that the disease has spread in such a way that the patient's cancer is incurable. The doctor believes that the surgeon has explained the situation to the patient and is going to express his sympathies to the patient. However, the surgeon has only told the patient that the tumor has successfully been removed and has not mentioned the metastases he found. The patient therefore incorrectly believes he is cured and is quite optimistic. When the doctor expresses his sympathies to the patient, the patient initially responds in a confused manner. After obtaining more information from the doctor about the situation the patient's conversational behaviors shift to behaviors characterized by displays of Shock and later by displays of Anger and Sadness. The conversation ends when the patient turns away from the doctor denying his situation. The doctor is unsure of what to say or do next and leaves the patient after saying he will return later.

This scenario gives a good illustration of the changes occurring in the patient's internal state as he hears the bad news and tries to deal with it. These changes to the patient's internal state features are reflected by the different expressions of emotion and the different types of conversational behaviors he performs. The course of the conversational behaviors also shows that the patient is adopting various ***coping strategies*** throughout the conversation.

1. At the beginning of the conversation the patient is unaware of the bad situation he is in. As a result he displays Joy and Relief emotions.

[1] For more information see www.pallium.ca

2. When the doctor tells the patient the bad news, we see the patient adopting a coping strategy of ***mild denial*** as described by Kübler-Ross in chapter [3]. The patient questions what the doctor is telling him as it contradicts his current beliefs.
3. When the doctor continues to tell the patient that he is in a bad situation, the patient adopts the ***anger*** coping strategy (also described in chapter [3]), while continuing to deny the doctor's words by questioning what he is saying. This coping strategy can be ascribed based on the observation of displays of Anger. Alternatively, the displays of Anger can also be ascribed to the strategy of ***venting emotions*** as described by the COPE classification of coping strategies. (See section [4.3.4])
4. After the expression of his anger, the patient adopts a coping strategy of ***strong denial***, refuting everything the doctor is saying and even offering alternatives.

In addition to the patient's coping strategies, the bad news conversation in the video-clip provides some insight into the features that hold in the doctor's internal state during the conversation. The doctor's conversational behaviors also indicate that the doctor fails to follow any of the protocols described in chapter [3]. One of the particular failings of the doctor's approach is that he does not establish what the patient knows and believes.

As mentioned earlier, two analyses of the conversational behaviors performed in the bad news conversation shown in the video-clip have been carried out. In this chapter we explain how these two analyses were performed and present and discuss the findings. In section [5.1], an overview of the assumptions about the relation between conversational behavior and internal state features is presented. Section [5.2] discusses the detailed observation and annotation that has been done on the bad news conversation performed in the video-clip in an attempt to label the displays and properties of the conversational behavior. Section [5.3] describes an online questionnaire that has been conducted in order to gain more insight in how the internal state of the interlocutors can be associated to the conversational behavior they perform and how such behavior is being perceived by people without a medical background. Conclusions that are drawn from the findings of both analyses are presented in section [5.4].

5.1 Relations between conversational behavior and internal state features

One of the assumptions made in this chapter, is that certain displays in conversational behaviors can be associated with specific internal state features that hold at the moment the behavior is performed. For example, when a person performs behavior that contains a display of grief, such as crying, it is reasonably plausible to assume that that person is experiencing feelings of Sadness. In other words, in the internal state of the person, the feature "emotion: Sadness" is present and active during the performance of the conversational behavior.

Furthermore, we argue that conversational behaviors have properties that implicitly express certain internal state features of the speaker. The internal state features

that are expressed by the `behavior properties` represent the motivations and expectations of the speaker. For example, a conversational behavior that contains the utterance "The surgeon has removed the tumor" has the `intended effect` (one of its properties) that expresses the speaker's **intention** (an internal state feature), namely to inform the listener about the fact that the tumor has been removed. At the same time the speaker may has the **expectation** (an internal state feature) that the listener performs a joyful response behavior. The speaker's expectation is expressed by the `expected effect` property of his conversational behavior.

By identifying the properties of a conversational behavior, we can make assumptions about which internal state features are related to that conversational behavior and thus which internal state features hold while the speaker is performing that conversational behavior. In addition, by identifying the purposes and effects of multiple conversational behaviors assumptions can be made about what ***strategy of coping*** the speaker is following.

As explained in more detail in section [2.4.1], `the intended effect` (or `Communicative function`) property of a conversational behavior can, in part, be identified by interpreting the semantic content of the conversational behavior. Recognizing certain keywords in an utterance can give an indication of the purpose of the conversational behavior. For example, if the utterance of a conversational behavior starts with words like "how", "why" or "where" this is a good indication that the `intended effect` of a conversational behavior is to *`gather information`*.

Note that utterances in conversational behaviors that have the intended effect of gathering information do not necessarily have to be formed as questions. For example, the intended effect of a conversational behavior that contains the utterance "Tell me the results of the test." is also to get more information.

After identifying the intended effect of a conversational behavior we can make assumptions about the intention, i.e. a feature of the internal state, which we believe is associated with the intended effect. For example, if the `intended effect` that can be identified from a conversational behavior is *`gather information`*, the **intention** that can be ascribed to the (internal state of the) speaker is that he wants (i.e. has the intention) to 'obtain a specific piece of information' from his interlocutor. One can assume that the intended effect is associated with the internal state feature **'Intention: *get information*'**.

Another property of a conversational behavior is its `type`. The type of a conversational behavior is used to describe the form of the conversational behavior regardless of its function. For example, "Am I getting better?" has the function of getting information, and the form, and thus the `behavior type`, of a *`question`*. An alternative type for the same function might be an *`order`*: "Tell me if I'm getting better." Behavior types say something about how conversational behaviors relate to other conversational behaviors. For example, if the type of a conversational behavior is a *`question`*, it is often followed by an *`answer`* type. It is often preceded by a *`statement`* type (performed by the interlocutor receiving the question) of conversational behavior, of which the speaker did not understand the stated information.

Convenient features to identify the types of conversational behavior are adjacency pairs. One might easily identify a conversational behavior to have the type

acceptance or *rejection* if the type of the preceding conversational behavior performed by the interlocutor is identified as an *offer*. In this thesis we use the following list of behavior types to distinguish the various conversational behaviors in the analyzed conversation:

Behavior types:

- *questions*
- *answers*
- *statements*
- *orders / suggestions*
- *expressives (greeting, apologies and condoling)*
- *acceptances / rejections*
- *excuses*
- *acknowledgements*

Interpretation and processing (i.e. evaluation with respect to the current features in the listener's internal state) of the speaker's conversational behavior is required before the properties of the conversational behaviors can provide information about the internal state features of the speaker. We are specifically interested in the properties that indicate which specific features of the internal state of the speaker hold while he is performing the conversational behavior.

The question is how to identify the properties of a conversational behavior. Often the properties can be determined most easily by considering various typical indicators. These indicators consist of facial expressions (most often to identify expressions of emotions), prosody (e.g. a rise in pitch at the end of a sentence often indicates a question), syntactic construction, (e.g. questions are structured differently than answers or statements), the behavior types of preceding conversational behaviors and the semantic content of the utterance.

In complex conversations it is almost always the case that multiple features of the internal state are expressed in a single conversational behavior through a variety of displays and conversational behavior properties. Making correct assumptions about the internal state features that gave rise to the conversational behavior can become difficult when the conversational behavior displays multiple features and has a range of properties. This is because sometimes interpretation of certain displays or properties may contradict the interpretation of other displays or properties, For example, if the speaker is using a loud tone of voice and uttering short explicit sentences, a listener could derive from this that the speaker's internal state contains a high intensity for the emotion **Anger**. If at the time the listener interprets the `communicative function` of the conversational behavior as trying to comfort the listener, it is unlikely that the underlying emotion of that conversational behavior is **Anger**. Ergo, the two internal state features ascribed to the speaker conflict with each other. Such a situation might occur when an utterance such as: "Stop crying! Everything is alright!" is performed in a raised voice (i.e. which is characteristic for Anger).

While the presence of multiple displays or behavior properties might hinder making assumptions about the internal state features underlying a conversational behavior, it is also possible that they make it easier to make the correct assumptions. For

example, when a person is crying, the most common interpretation of the behavior is that the person is displaying the emotion Sadness. However, if the person utters the phrase "I'm so relieved" at the same time, the internal state feature underlying the crying behavior can be interpreted as **Happiness**, i.e. tears of joy.

Another difficulty is that the observed displays or interpreted properties of a conversational behavior do not necessarily correspond to or represent the beliefs and intentions that hold in the internal state of the speaker at that moment. For example, the conversational behavior of a parent who is scolding a child will contain displays of Anger, even if the parent is not experiencing this anger internally. However, the child's interpretation of the conversational behavior would cause him to ascribe the emotion Anger (i.e. an internal state feature) to the speaker (i.e. the parent). In other words, the child thinks the parent is angry.

Two analyses of the conversational behaviors performed in the Pallium video-clip were subsequently carried out. The goal of these analyses was to make more assumptions about the relations between conversational behaviors and underlying internal state features, and to strengthen the assumptions we already made. During the process of making more assumptions about the internal state features, we paid attention not to make the possible mistakes described in the previous paragraphs.

5.2 Analysis of observed conversational behavior

As information about the relation between a person's internal state features and the conversational behaviors he performs during a bad news conversation is sparsely covered in the literature, we decided to gather additional data on these relations. To this end, we performed an analysis of a set of conversational behaviors that were used in a bad news conversation. More specifically, we analyzed the conversational behaviors of the bad news conversation that was enacted in the first part of the Pallium video-clip. The purpose of this analysis was to gain insight in what kind of displays and properties can be identified in conversational behavior performed in a bad news conversation. The second purpose of the analysis was to see if the assumptions we make about the relations between **internal state features** and the identified `displays and properties` of the conversational behaviors can be supported. We argue that a significant number of the internal state features that hold can be determined by analyzing the performed conversational behaviors. The analysis was done by closely observing and annotating the displays and properties of the conversational behaviors performed in the Pallium video-clip.

5.2.1 Annotation scheme

The annotation was done on the basis of a rough annotation scheme we created. Instead of consisting only of a strict set of instructions (i.e. labels), the annotation scheme also contained a set of guidelines (i.e. pointers) that could by the participants during the annotation process. This was done because it was expected that observation of the conversational behaviors would also yield findings that could not easily be labeled.

While the labels were specific instructions that were quite straight forward on how to describe certain displays, the pointers were meant to indicate which features of the conversational behavior an annotator should pay attention to when describing which displays and properties could not be described by the labels.

Examples of such pointers are: "look at where the doctor places his hands" or "mention how close the doctor is standing to the patient" or "The following words are quite likely to be expressions of the matched internal state features: I think belief, I want goal/desire, I will be belief (expectation)."

The annotation scheme consisted of eight categories in which labels and pointers were given that could be used to describe the displays observed in the conversational behaviors and the properties ascribed to the behaviors. The eight categories containing the labels and pointers are listed below:

1. The emotions displayed through facial expressions.
2. The emotions displayed through prosodic information.
3. Other internal state features expressed through prosody.
4. The interpersonal relation displayed through body posture.
5. Internal state features expressed through hand and/or head gestures.
6. Internal state features expressed through explicit statements.
7. The Behavior Type or form of the uttered sentence.
8. The Dialogue Act Type of the uttered sentence.

In the rest of this section we discuss these categories in more detail and describe the labels and pointers which they contain. The labels and pointers present in these categories were based on the models and theories presented in previous chapters. From chapter [4] we use the information about possible ways in which features of the internal state can be expressed. From chapter [2] we use information about the dialogue act types of the conversational behavior (i.e. `the communicative function` or `intended effect`), to determine the labels and pointers. Furthermore, intuitive notions about how we think people interpret displays and properties in real life conversations were taken into account.

Catergory 1:

The first category, **Emotions through Facial Expressions**, contains labels and pointers that can be used to describe the displays of possible emotions the interlocutor might be experiencing. As mentioned in chapter [4], we use the **22 emotion types** described in the OCC model of emotions developed by Ortony, Clore and Collins (Ortony et al., 1988) to represent the emotions that are present in a person's emotion state. To this we added four additional emotion types of our own to represent the internal states **Surprise, Confusion, Sympathy** and **Empathy**.

However, not all emotion types can be expressed through facial expressions, as a display of emotion. For example, the emotion type Hope does not have a distinctive facial expression that can be related to it. Furthermore, some of the emotion types in the OCC model appear to be closely related to others, making it difficult to distinguish between the facial expressions that convey to them. (Provided there are facial expressions that can convey them). This holds for emotion types such as **Gratitude** and d7**Gratification**, **Sadness** and **Pity** (i.e. **Sorry-for**), and **Joy** and **Happy-for**.

In order to annotate the displays of emotion through facial expressions that are recognizable, we use the list of Ekman (1971) which contains six basic emotions (i.e. Joy, Sadness, Anger, Fear, Surprise and Disgust) as labels. However, we included pointers into the category to also be able to label facial expressions as displays of the following emotion types: Happy-for, Pity, Confusion and Relief. These pointers suggest to the annotator to take into account ...

- who the subject of the displayed emotion is. Joy, Sadness, Fear, Surprise and Relief are emotions relating to the speaker, while Anger, Disgust, Happy-for and Pity are directed at or relate to someone (or something) other than the speaker. This pointer is aimed at allowing the annotator to make a distinction between emotion types that are closely related to each other.
- how the displays of emotion relate to previous conversational behaviors. This can be done by making assumptions about the internal state features of the speaker that hold prior to him performing a facial expression to display an emotion.

Although the observed facial expressions that displayed the emotions present in the internal state, were often indicated by the provided labels, the labels were sometimes extended with descriptions. These additional descriptions were based on the lists of tokens given for each emotion type in Ortony et al. (1988). (see appendix [B] for a more detailed overview of the emotion types). These extra descriptions provided additional information about how the various emotion types could be expressed. For example, observation of a display of Anger by the patient could be described as "the patient is looking annoyed."

Catergory 2:

Similar to the first category, the second category, **Emotions through Prosody**, contains labels that are used to describe expressions of the speaker's emotions. However, the prosody used in utterances performed by a speaker can cover a wider range of emotion displays than facial expressions. More specifically, a speaker can use prosody to express all emotion types that are part of the speaker's emotion state. Consequently, the second category uses all **22 emotion types** described in the OCC-model as labels to describe the displays of emotion expressed through prosody plus the **additional four emotion types** we devised ourselves. For instance, a description of such a display could be "the doctor sounds sad." Pointers included in the category point out characteristics of the prosody that the annotator should pay attention to. These characteristics include:

- the rise and fall of the pitch.
- the loudness (i.e. energy) of the utterance.
- the rate of speech.
- the tone of voice (creaky, hoarse, throaty, whispery etc).

Specific combinations of the values of the characteristics can indicate certain emotions. For example, **Anger** is characterized by a lower pitch, a higher loudness (i.e. intensity) and a faster onset of speech in comparison with neutral speech. Compared

to neutral speech, **Sad** speech is characterized by a higher pitch, a lower loudness and slower rate of speech, combined with longer pauses between words.

Catergory 3:

In addition to providing information about possible emotions the speaker might be experiencing, the prosody of an utterance can also display other features of the speaker's internal state. Descriptions of such displays are grouped in the third category, **Other Internal State Features through Prosody**. These other features are the speaker's *mood* (e.g. cheerful or depressed), *affect disposition* (e.g. nervous, anxious, irritated, reluctant or envious) or his view of the *interpersonal relation* (also known as the interpersonal stance) that exists between him and his interlocutor (Scherer, 2005). For example, disfluency in speech might be an indication of nervousness or reluctance of the speaker.

Because the annotation scheme was constructed before the detailed analysis of the video-clip, it was not quite clear what kinds of prosodic information would be present in the conversational behaviors. As a result, no specific labels were created or selected beforehand for this category but only pointers were included. The pointers called the annotator's attention to the quality of the voice and the fluidity of the speech. Similar to the pointers in category two, the annotator is suggested to focus on:

- the rise and fall of the pitch.
- the loudness (i.e. energy) of the utterance.
- the rate of speech.
- the tone of voice (creaky, hoarse, throaty, whispery etc).

Catergory 4:

The fourth category, **Interpersonal Relation through Body Postures**, covers descriptions of the body postures of the interlocutors and assumptions about their dispositions towards each other. Although this category does not include predefined labels, it does offer pointers that could be taken into consideration when describing and interpreting the body postures. The focus of describing and subsequently interpreting the body postures lays on the openness and reticence (i.e. the degree of reserve) of the posture and the proxemics of the interlocutors. The pointers for this category are based on the concept of interpersonal stance and direct the attention of the annotator to:

- the distance between the two interlocutors.
- the body and facial orientation of both the speaker and the listener.

For more information about interpersonal stances with respect to body postures see Bickmore (2008) and Scherer (2005). Specific configurations of the body posture and the proxemics may be interpreted as displays of the interpersonal relation according to the person observed. For example, a description such as "the doctor keeps quite a distance from the patient. He has his arms low with his hands clasped in front of his body." might be interpreted as the doctor taking a reserved stance towards his interlocutor.

Catergory 5:

The fifth category, **Internal State Features through Gestures**, focuses on describing displays of internal state features that are composed of gestures of the hands and movements of the head. The labels contained in this category describe the function that may be expressed by the head or hand gesture. The following labels are included in this category:

- Head nod expressing a backchannel.
- Head nod expressing confirmation or agreement.
- Head shake expressing disconfirmation or disagreement.
- Pointing gesture with the head or hand to direct attention to something.

In addition to the labels, the following pointer was provided in category five to assist the annotator in describing displays expressed through head or hand gestures.

- Pay attention to how the speaker's gestures have an affect on his interlocutor.

For example, when the doctor places his hand on the shoulder of the patient, the patient calms down. Such a gesture might be interpreted as a display of the doctor's intention to comfort the patient.

Catergory 6:

Category six, **Internal State Features through Explicit Statements**, includes labels that describe speech utterances that express features of the internal state that hold while performing that utterance. Note that not all conversational behaviors contain a speech component that explicitly articulate the speaker's internal state features, but many of them do. Category six contains labels to describe expressions of all the different internal state features that were discussed in chapter [4]:

- The **beliefs** of the speaker (including his expectations).
- The **goals** of the speaker.
- The **intentions** of the speaker
- All **22 emotion types** from the OCC model
- The emotions **Confusion, Surprise, Sympathy** and **Empathy**
- The social features of **rapport, liking** and **trust**

The expression of internal state features is more explicit in some speech utterances than in others, depending on which words the speaker is using. For example, when the patient speaks the utterance "But I thought the surgeon removed the tumor." it is quite obvious that he previously had, (and perhaps still has) an internal state of **belief** (i.e. his thought) that the surgeon removed the tumor. Another example is the sentence "I'm afraid that the chemo will not be sufficient to cure the cancer." which expresses the doctor's emotional state of **Fear**. However, in a sentence such as "Hasn't the surgeon talked with you about the results?" the speaker's internal state features are expressed a lot less explicitly. In such cases where expression of the speaker's internal state features is more or less implicit, the verbal content of the speech utterance first needs to be interpreted before the expressions can be labeled.

For example, based on an interpretation of the utterance with implicit expressions of the internal state, the following assumptions can be made: The speaker previously had the belief that the surgeon had talked to the patient, but now the certainty of that belief has decreased. Although the category does not have strict guidelines that indicate how utterances should be labeled, the following pointers are suggested to the annotator:

- Utterances that contain phrases such as 'believe', 'think', 'suspect' and 'know' may be tagged with the belief label as these utterances contain expressions about the speaker's internal state features **beliefs**.
- Utterances that contain phrases such as 'want' and 'would like' to may be tagged with the goal label as these utterances contain expressions about the speaker's internal state features **goals**.
- Utterances that explicitly mention one of the **26 emotion types** (i.e. the 22 of the OCC model, plus the additional four) or one of its tokens (see appendix [B]) may be tagged with the corresponding emotion label, if the expression concerns the speaker. For example, "I'm sad that the cancer cannot be cured."

Catergory 7:

The seventh category, **Behavior Type**, contains labels that are used to describe the form of the conversational behavior performed. We consider the behavior type to be a property of a conversational behavior, rather than a display or expression of a feature of the speaker's internal state used in the conversational behavior. As such, we cannot use behavior types to make assumptions about the internal state features that produce the conversational behavior. However, `behavior types` are useful in determining how a conversational behavior may relate to other conversational behaviors. Combining the behavior type with rules for adjacency pairs may help determine what type of behavior is appropriate for the listener to respond with. Labels that are contained in this category are *`questions, answers, statements, orders/suggestions`* and *`expressives`* (Expressives include behaviors such as greetings, apologies and condoling. See section [2.3] for more information on expressives). Pointers that are included in this category suggest to the annotator to look at the `behavior type` of the preceding conversational behaviors.

While they are both properties of a conversational behavior, `Behavior types` differ from `dialogue act types` in the sense that behavior types describe the *form* of the conversational behavior rather than categorizing the *function* of the conversational behavior. For example, the sentence "What do you think about the diagnosis?" has the same communicative function as the sentence "Tell me what you think about the diagnosis.", i.e. to *`gather information`* about the beliefs of the interlocutor. However, the first sentence has the form of a *`question`*, while the second sentence has the form of an *`order`*. Another example are the sentences "Can you hand me my medicine?" and "What did the doctor say?" Both sentences have the behavior type of a *`question`*, but the intended effect (i.e. dialogue act type) of the first sentence is to make the interlocutor *`perform an action`* (rather than provide the answer "Yes, I can."), while the intended effect of the second sentence is to *`gather information`*.

Catergory 8:

Included in the eighth category, **Dialogue Act Type**, are labels that can be used to indicate what the type of dialogue act (i.e. the conversational behavior) is. As discussed in section [2.4] in the second chapter, in this thesis we use the term `dialogue act type` to indicate the `communicative function` of a conversational behavior (i.e. the `intended effect` of a conversational behavior). This is similar to the approach Bunt takes when discussing the DIT++ taxonomy (Bunt, 2009). As a result, the labels that are included in this category are a selection of the dialogue act types described in the DIT++ taxonomy. The selection of labels consists of the following dialogue act types:

`Dialogue act types:`

- *question*
- *inform*
  - *agree (specific type of inform)*
  - *disagree (specific type of inform)*
  - *confirm (specific type of inform)*
  - *disconfirm (specific type of inform)*
- *instruct (i.e. order)*
- *request*
- *suggest*
- *auto-feedback (see chapter [2] for more information)*
- *allo-feedback*
- *self correction*
- *open conversation*
- *close conversation*
- *introduce / Shift topic*
- *salutation (social relation function)*
- *self introduction (social relation function)*
- *apologize (social relation function)*
- *gratitude expression (social relation function)*
- *valediction (social relation function)*

The `intended effect` property of a conversational behavior can be related to the internal state feature **intention** that holds while the behavior is performed. Though not an actual display, the intended effect expresses the intention implicitly through the conversational behavior's semantic content. Similar to the behavior types, we consider the dialogue act type (or more specifically the communicative function or intended effect) to be a property of the conversational behavior. Determining what the intended effect of the conversational behavior is can be a difficult task. The pointer that is included in this category suggests to the annotator to think about:

- what behavior type label the (preceding) conversational behavior has.
- what goals might the speaker have in this situation.

For example, often when a person asks a question, the intended effect of the conversational behavior is to *receive information* from his interlocutor by getting an *answer* as a response.

5.2.2 Annotation and findings

In order to be able to make grounded assumptions about the relations between the internal state features of the interlocutor and the conversational behaviors he performs, we needed to analyze the behaviors. To this end we annotated the conversational behaviors that were performed in the Pallium video-clip. This was done using the annotation scheme presented in the previous section, which was supported by the insights we gained from the theories and models discussed in chapters [2], [3] and [4]. As this is a precursory analysis of the conversational behaviors and an initial application of the annotation scheme, we performed the annotation ourselves to see how the annotation scheme would operate.

To facilitate the analysis, the video-clip was divided into smaller fragments with each fragment consisting of the conversational behaviors performed in a single dialogue turn. All conversational behaviors per dialogue turn were analyzed, not just those performed by the speaker, but those performed by the listener as well. After observing the conversational behaviors in each dialogue turn, we decided to group some of the dialogue turns together, based on similarities in the findings and on their function in the conversation. Overall, we defined four segments in the video-clip. In the next paragraphs we give a description of the context of each segment and present transcripts of the speech utterances.

In addition to the context descriptions and the transcripts, we discuss the most interesting displays of internal state features and properties of conversational behaviors performed in each separate dialogue turn (i.e. fragment) as well as displays that stretch out over multiple dialogue turns.

Segment 1:

In the first segment, the doctor and the patient introduce themselves to the viewers of the video clip. In these dialogue turns (i.e. D1 and P1) both interlocutors describe the situation from their perspectives and present the context in which the bad news is placed. Even though these two dialogue turns are not actually part of the bad news conversation, they provide the viewer with insight in some of the thoughts and feelings the interlocutors have with respect to the situation before engaging in the dialogue. Therefore we included these dialogue turns in the annotation. Segment one also contains the dialogue turns that cover the opening of the conversation through an exchange of greeting behaviors by both interlocutors and the expression of the doctor's feeling of empathy with regard to the patient's situation. Segment one consists of five dialogue turns, containing the following speech utterances:

Doctor (D1): *I'm about to visit John Filpot. He is recovering from surgery. He had a bowel tumor removed and evidence of metastases. This means that John's cancer is incurable. Now, I know the surgeon has talked with John, but I don't know if John really understands his condition.*

Patient (P1): *I'm feeling pretty good. Yeah I got some pain, but the surgeon got my tumor. It was attached to my bowel. But the most important thing is that the tumor is gone. I'll probably be out of here in a few days, once they can remove these tubes and get me on some decent food. I'm sure everything is going to be okay. No biggy, as my son says.*

At this point, the doctor enters the patient's room and the actual bad news conversation between them starts.

Figure 5.1: Screenshot from the Pallium video clip

D2: *Hey John.*
P2: *Hello doctor Rupert.*
D3: *I'm sorry to hear the news. We were hoping that surgery and a course of chemo would cure this tumor.*

In the first segment, a wide variety of displays and properties of the conversational behaviors was annotated. In dialogue turn D1 (the introduction of the doctor), the doctor is using short sentences and there are long pauses between each sentence. In addition he uses a soft voice (i.e. low energy). This combination of displays indicates that the doctor does not want to speak to the patient and subsequently the behavior is annotated as the speaker's affect disposition of being *reluctant*. The description given to these displays falls into the third category of the annotation scheme. Furthermore, the low pitch of the voice combined with the low energy and the slow rate of speaking, are indications of the emotion **Sadness**. Subsequently, the utterance is given the label Sadness from the second category.

The facial expression of the doctor in D1 shows the doctor is feeling uncomfortable. This is one of the tokens of **Distress** (i.e. **Sadness** in Ekman's list) described in the OCC model. Interpreting the facial expression as a display of Sadness is done based on the observation of the facial features of lowered mouth corners and raised inner eyebrows. The facial expression is described with the label Sadness from the

first category. The displays of Sadness expressed through both the doctor's use of prosody and his facial expression are maintained throughout the dialogue turns that make up the first segment and reflect his emotion state in this part of the conversation.

Based on the semantic content of the utterances performed in D1 we can derive additional expressions of the doctor's internal state features. Particularly, the last sentence in D1 contains explicit expressions of the doctor's **beliefs**, namely "I know ..." and "I don't know ...". Consequently, we labeled these sentences with the label **belief** from the sixth category. The sentences in dialogue turn D1 are statements of facts about the situation and receive the *statement* label for their behavior types. Based on this, we assume that these statements express the doctor's belief that these facts are true.

Based on the context of dialogue turn D1 (i.e. the doctor informs the viewer of the instructional video-clip about the situation before the bad news conversation starts) and the behavior types used in dialogue turn D1 (i.e. *statements* of facts), we argue that the communicative functions of the utterances are to *provide the viewer with information* (i.e. the conversational behaviors are given the label *inform* from the eighth category to describe their dialogue act type).

In dialogue turn P1 (the introduction of the patient), the prosody of the patient's utterances is characterized by a high pitch and high energy level. Combined with a neutral rate of speech, we interpret the prosody as a display of a *cheerful* affect disposition (a display of an "other internal state") and a moderate *display of Joy* (a display of emotion). The patient's facial expressions in dialogue turn P1 are displays of Joy (or rather the token Contentment). These are recognizable by the raised mouth corners and the slightly raised cheeks. This display of emotion through the patient's facial expression is labeled as Joy.

The semantic content of the speech utterances indicates that the patient is expressing his **beliefs** about the situation, as they contain *statements* of facts: "the surgeon got my tumor." and "The tumor is gone." In addition, the utterances express one of the patient's **expectations**, namely that he is going home soon: "I'll probably be out of here in a few days". Based on the **beliefs** we associate with these speech utterances we draw the conclusion that the patient incorrectly believes that everything is alright. In the introduction of the doctor we have observed that this is not the case. The dialogue act types of the conversational behaviors in dialogue turn P1 are labeled with the *inform* label from category eight.

The greeting and the return greeting behaviors in dialogue turns D2 and P2 contain few displays that can be described by labels from categories one to seven. The conversational behaviors in these dialogue turns are only labeled for their `communicative function` (i.e. dialogue act type). The behaviors in both D2 and P2 are labeled as having a *salutation* dialogue act type. This indicates that both interlocutors treat each other politely by adhering to social obligations.

Dialogue turn D3 contains a number of displays that express the doctor's emotions. The prosody of the doctor's utterances has the characteristics of a *display of Empathy*: low pitch, low energy and an even intonation and was labeled with the Sorry-for emotion from the second category. The interpretation of this display is strengthened by the semantic content of the speech utterances performed in D3. In

the utterance "I'm sorry to hear the news." the doctor explicitly expresses his emotion of **Empathy** towards the patient. This display is assigned the label Sorry-for taken from the sixth category. For these displays we deduce that the `intended effects` (i.e. communicative functions) of the conversational behaviors are to *convey* the doctor's *sympathy and empathy* to the patient. In other words, the dialogue act type label given to this utterance is *inform*. (To inform the patient about the doctor's emotion state.)

Based on the annotations of the conversational behaviors in dialogue turn D3, we make the assumption that the empathy expressed by the doctor is a consequence of his **belief** that the patient is experiencing negative emotions. In turn, this belief is caused by another belief of the doctor, namely that the patient is aware of the bad news situation. We know that this last belief holds in the doctor's internal state as he expresses this in dialogue turn D1: "I know the surgeon has talked with John." As the patient has not had the opportunity to express himself in the conversation besides greeting the doctor, the beliefs the doctor has about the thoughts and feelings of patient are based on assumptions the doctor has made beforehand. As mentioned in chapter [3], the recommended approach for the doctor would have been to ask the patient about his thoughts and feelings before performing the conversational behaviors based on his (the doctor's) assumptions. In that case, the doctor's conversational behavior of expressing Empathy towards the patient could have been adjusted to something more appropriate. This would have served both the doctor and the patient better, but now the utterances of the doctor cause the patient to become confused.

This confusion is caused because what the doctor is saying contradicts with what the patient believes. As a result of the confusion, the patient questions the doctor's words. This approach can be construed as a denial strategy of coping with situation as explained in chapter [3].

Segment 2:

The conversational behaviors in the second segment reflect the patient's **Confusion** about his situation, his **intention** to gain more information from the doctor and subsequently the conversational responses of the doctor to the patient's inquiries. Segment two consists of the following speech utterances performed by the patient and the doctor:

P3: *What do you mean?*
D4: *When a tumor has spread through the abdomen, like yours has, there is a limited amount we can do for you.*
P4: *What do you mean, spread through the abdomen like mine has? The surgeon said he had removed the tumor.*
D5: *Yes, The primary one. But he couldn't get all the places it had invaded, we need to give you chemo for that.*
P5: *What do you mean invaded? You keep talking about this chemo-stuff.*

Throughout the patient's dialogue turns in segment two (i.e. P3, P4 and P5) displays of the patient's confusion are present in his conversational behaviors. An-

notation of the conversational behaviors in segment two resulted in the labeling of facial expressions with Confusion labels. *Displays of Confusion* through facial expressions are characterized by a creased forehead, eyebrows pulled together and a lowering of the jaw. In addition to making facial expressions to display his confusion, the patient also expresses his emotion via the prosody of his speech utterances. The prosody is labeled confusion as well. This second confusion label is part of the second category, whereas the first confusion label is part of the first category. Further observation reveals that the patient's speech has an alarmed undertone which may be an indication that even though the patient does not know what is going on, he **believes** that something is not well.

Figure 5.2: Screenshot from the Pallium video clip

In addition to displaying the emotion through his facial expressions and the prosody of his speech utterances, the patient's confusion is also expressed through the types of his behaviors. The conversational behaviors in the patient's dialogue turns (i.e. P3, P4 and P5) all start with speech utterances that are labeled with the conversational behavior type *question*, which falls into the seventh category.

Utterances with the *question* behavior type almost always have the `intended effect` to *get information* from the interlocutor. The dialogue act types of these utterances are therefore labeled as a question, with the label taken from the eighth category. We make the assumption that, because the semantic content of the speech utterances in D3 and D4 contradicts the patient's old beliefs about the situation, the patient is using the *question* behavior types and dialogue act types of *gather information* about the situation to verify his **beliefs** and subsequently resolve his confusion.

The second parts of dialogue turns P4 and P5 contain explanations why the patient is asking these questions, i.e. "The surgeon said he had removed the tumor." and "You keep talking about this chemo-stuff." Based on these explanations we make the assumption that the patient is giving the doctor additional information as to why he is

asking the questions. More specifically, the second utterances in P4 and P5 *convey* the patient's *disagreement* with what the doctor is saying, by indicating that the doctor's information contradicts with what the surgeon told the patient. Consequently, the second utterances in dialogue turns P4 and P5 were labeled as disagree dialogue act types.

Based on these annotations, we conclude that the patient is trying to cope with the situation by holding on to his old **beliefs**. This approach to the situation is part of the ***Denial*** coping strategy described in chapter [3]. The patient's old **beliefs** are positive with respect to the situation, opposed to the new information provided by the doctor. However, at the same time the patient tries to get more information from the doctor by posing questions, which indicates that he might be questioning the certainty of his old beliefs.

Although the conversational behaviors performed by the doctor in this segment are primarily aimed at providing the patient with (additional) information about the situation, they contain additional displays and properties that express features from the doctor's internal state. In the first part of the conversational behavior in dialogue turn D4, the doctor displays his Confusion through his facial expression as well as through the prosody of his speech utterance.

Figure 5.3: Screenshot from the Pallium video clip

The *display of Confusion* through prosody is characterized by a hesitating pause at the start of the doctor's speech utterance, which may indicate that the doctor is unprepared on how to respond to the utterance performed by the patient in P3. The confusion labels are taken respectively from the first and the second category.

The last part of the speech utterance performed in dialogue turn D4 contains displays of the emotion Sympathy instead of Confusion, which is expressed through the doctor's facial expression. This shift in emotions is a good indication that the doctor is experiencing multiple emotions in at the same time, but that only one of the emotions can be expressed at a particular moment.

The prosody of the speech utterance performed in dialogue turn D5 expresses the doctor's *annoyance* towards the patient. The behavior displaying the doctor's annoyance is characterized by stronger vocal expression of the words in the utterance, as well as faster rate of speech (i.e. faster pronunciation of the individual words). The prosody of the doctor's annoyance is labeled Anger and falls into the second category. We make the assumption that this display of Annoyance is the result of the doctor being displeased about having to explain the situation to the patient (as he believes the patient is already aware of the bad news situation). What is interesting is that the doctor's facial expression does not match the annoyance that is expressed by his voice. Instead the doctor returns to displaying a rather moderate expression of confusion about the whole situation.

The body posture the doctor is having during the conversational behaviors in this section is as following: the distance he is standing from the patient can be described as normal or typical for a formal conversation. However the doctor has his hands crossed in front of his body, signaling a form of defensiveness. This may stem from his belief that the conversation is not going to be pleasant.

The speech utterances in dialogue turns D4 and D5 were labeled with the *answer* `behavior type` label from the seventh category. This is because the semantic content and the form of the utterances in both dialogue turns were determined to be appropriate responses to the *question* `type of behavior` performed in the dialogue turns preceding them, respectively P3 and P4.

By using the assigned behavior types and the semantic content, we labeled the `communicative functions` of the conversational behaviors in dialogue turns D4 and D5 as ways to *inform* the patient about the correct state of the situation, i.e. to deliver the bad news to him. These annotations fall into category eight.

Based on the annotations we assigned to the conversational behaviors in the second segment, we draw the following conclusions. The doctor's confusion is caused by the patient's unexpected (for the doctor) response to the expression of Sympathy in dialogue turn D3, namely posing a question, instead of expressing gratitude for the doctor's display of Sympathy or at least acknowledge it. As a result of the patient posing a question about the situation, the doctor may become uncertain about his beliefs about what the patient does and does not know. If the doctor followed a protocol such as **S.P.I.K.E.S.**, he would not experience Confusion about the patient's lack of knowledge.

Similar to his behaviors contained in the first segment, the patient continues to question the information the doctor is giving him. As a result he is effectively denying the doctor's bad news. As mentioned earlier, this can be construed as a denial strategy of coping.

Segment 3:

The dialogue turns in segment three contain conversational behaviors that are aimed at resolving the confusion that has arisen between the patient and the doctor by finding an explanation for the misunderstanding. Furthermore, the conversational behaviors focus on clarifying the state of the situation for both interlocutors. The transcript of the speech utterances performed in the dialogue turns is as follows:

D6: *Didn't the surgeon talk with you?*
P6: *Yeah but he didn't say anything about chemo or invasive stuff. What the hell is going on here? I thought I was cured, but now it sounds like I'm not.*
D7: *I'm sorry John. I thought you knew.*
P7: *Well I didn't! I don't believe it. (Pause) They have shown you the wrong chart! I'm going to be fine (camera shift). Just need a little medicine and I'll be right as rain.*

The last three utterances of dialogue turn P7, i.e. the part after the pause in the patient's speech, were not considered to be part of segment three. In our opinion, these utterances were more suited to be included in the fourth segment, based on their function in the conversation. They are only included in the transcript above to maintain the continuation of the dialogue. In the other utterances in segment three (i.e. the utterances in dialogue turns D6, P6, D7 and the first two utterances of P7), we again observed interesting displays and conversational properties of the behaviors.

The conversational behavior in dialogue turn D6 contains a moderate display of the doctor's **Confusion** expressed through his facial expression, which is carried over from dialogue turn D5. Similar to the conversational behavior in dialogue turn D4, this facial expression of Confusion is strengthened by expressions of confusion via the prosody of the doctor's speech utterances. In addition, a change in the doctor's body posture was observed. The doctor leans in towards the patient as an act of comforting and an indication of increased attention. We make the assumption that the increase of proximity is intended to increase the **rapport** between the patient and the doctor. Based on the form of the speech utterance, we assign the label *question* from category seven to the behavior type of the conversational behavior in dialogue turn D6. The `communicative function` of the conversational behavior is to get information from the patient, which we derive from the fact that the behavior type is a *question*. As a result, the dialogue act type is also labeled *question*.

The conversational behaviors in dialogue turn P6 contain strong displays of different emotions of the patient that follow each other in quick succession, expressed both in his facial expression and in the prosody of his speech utterances. During the first utterance in P6, the patient raises his voice, increases his rate of speech and uses clipped words. Although it sounds like the patient is accusing the doctor of causing the confusion about the situation (i.e. expressing **Reproach**), the observed prosodic characteristics can indicate displays of several different emotions, including **Distress, Anger** or **Reproach**. According to the OCC-model, the emotions Anger and Reproach are both brought about when (at least) the following situation holds: the person experiencing the emotion "disapproves someone else's blameworthy action". The emotion **Anger** has the additional premise that the person experiencing the emotion is "displeased about the related undesirable event." Based on the observed prosody and the information presented in the OCC-model, we argue that the patient is displaying his Anger rather than his Reproach towards the doctor. As a result we have labeled the display with the anger label from the second category. However, the distinction between whether Anger or Reproach is displayed is not very obvious in this case and one might argue that both emotions are expressed simultaneously.

The annotation of the display of Anger is supported by the observation of the expressions in the other two utterances of P6. The second and third utterances in P6 have a prosody similar to the first utterance, but in addition the speech is accompanied by facial expression that displays Anger. The anger label from the first category was assigned to the facial expression.

Figure 5.4: Screenshot from the Pallium video clip

In addition to the prosodic information and the facial expression, the patient also expresses his **Anger** through the semantic content of the utterances performed in dialogue turn P6. More specifically, the patient is using the expletive "What the hell . . . ", which is a good indication of the patient being displeased about the situation and **Blames** someone for it. The use of the expletive also reflects the patient's criticism about the situation (i.e. the way in which information is communicated to the patient by the doctor and the surgeon) and indicates that the social relationship between the doctor and the patient changes from *formal* to *informal* (there is a large decrease of politeness).

At the end of the third utterance in dialogue turn P6, the patient changes the expression of his emotions by displaying Sadness through his speech and face: his voice trembles slightly and there are tears forming in his eyes. Both displays are labeled with the Sadness label from the first and second categories.

The third utterance in dialogue turn D6 also contains a clear expression of one of the patient's internal state features, namely one of his **beliefs**. We argue that by using the phrase "I thought I was cured . . . ", the patient expresses how he **believes** (or previously **believed**) the bad news situation to be. Consequently, we labeled the utterance with a belief label from the sixth category.

The three speech utterances in dialogue turn P6 are given the following labels for their `behavior types`: The first utterance has the form of an *`answer`* and is given in response to the *`question`* posed by the doctor in dialogue turn D6. The behavior type of the second utterance is labeled as a *`question`*, while the third utterance has the

behavior type of a *statement* and is labeled as such. All labels and behavior types fall into category seven.

One of the interesting parts of dialogue turn P6 is the set of communicative functions that are expressed by the patient's conversational behavior. We argue that the conversational behavior in P6 has several intended effects that the patient would like to see fulfilled. Based on the abundant display of emotion in the conversational behavior performed in P6, it is very likely that one of the intended effects is that the patient wants to *inform* the doctor about how he is feeling emotionally. The expletive expresses the patient's Anger clearly, but also seems to be used to criticize the doctor for the **Confusion** the patient is experiencing. (Arguably expressing **Reproach** in addition to expressing Anger). The tears and the trembling voice indicate the patient's intended effect to *inform* the doctor about his Sadness. A second intended effect that we assigned to the conversational behavior in P6 is that the patient *confirms* the doctor's question whether the surgeon has spoken with him (i.e. the patient). Next, the patient expresses his *disagreement* with what the doctor is telling him about the situation, which is the third intended effect labeled. The intended effect of the third speech utterance in P6 may be to *inform* the doctor about the **beliefs** of the patient, in an attempt to clarify the doctor's Confusion.

The conversational behavior performed by the doctor in dialogue turn D7 is characterized by *displays of Self-reproach* (i.e. the emotion-type of the OCC-model which includes the tokens *embarrassment* and *shame*). The prosody of the speech utterances is characterized by a soft voice and a slow speaking rate, while the facial expression displays **Embarrassment**, i.e. pressed lips and lowered inner eyebrows. In addition, the doctor lowers his head and tries to avoid eye contact with the patient. These actions are also characteristic for expressing Embarrassment and might be placed into category five.

Based on the semantic content of the utterances, we argue that the intended effects of the conversational behaviors are the following. The intended effect of the first utterance is to comfort the patient by *informing* him about his (i.e. the doctor's) feeling of Sympathy, by saying "I'm sorry John." The same speech utterance, and the second utterance, may also have another intended effect, which is to *apologize / give an excuse* for bringing the information to the patient's attention: "I'm sorry John." (apologize) and "I thought you knew." (give excuse).

As mentioned earlier, only the conversational behaviors performed before the pause were considered to be part of segment three. The first two speech utterances in dialogue turn P7 contain displays of the patient's Sadness expressed in both the prosody and in the patient's facial expression. These displays are a continuation of the display of Sadness expressed at the end of dialogue turn P6.

The second utterance also contains a clear display of one of the patient's internal state features, namely "I don't believe it." The speech utterance indicates that the patient does not **believe** the information provided by the doctor and returns to his old **beliefs** about the situation that everything is okay. This behavior is a good example of the coping strategy of ***denial*** as described in chapter [3].

The behavior types of the two utterances are respectively labeled as an *answer* and a *statement* and fall into category seven. The first utterance has the commu-

nicative function to *disconfirm* the doctor's **belief** that the patient was aware of his situation, which he (i.e. the doctor) made clear through his preceding dialogue acts. It also disconfirms the doctor's previous statement "I thought you knew." The second sentence indicates the patient's *disagreement* with the doctor's beliefs.

Based on the annotations of the conversational behaviors performed in segment three, we draw several conclusions about the cognitive processing of the conversational behaviors by both interlocutors. The conversational behaviors performed by the doctor in segment three express that a shift in his beliefs has occurred. During the dialogue turns of the first and second segments, the doctor believed that the patient was aware of the bad situation he was in. However, the speech utterances performed by the patient in this segment have caused the doctor to form new **beliefs** (i.e. the patient is not aware of his situation), which contradict with his old beliefs (i.e. the patient is aware of his situation). This becomes especially clear in the second utterance of dialogue turn D7: "I thought you knew." In order to find out why the patient possessed incorrect beliefs about his situation, which were opposite to the old beliefs of the doctor about the beliefs of the patient, the doctor poses the question in D6. The speech utterance in D6 also implicitly explains why the doctor's previous speech utterances were contradicting the patient's beliefs, namely because the doctor assumed that the surgeon had explained the bad news situation to the patient. As a result, the relation between this speech utterance and the speech utterances performed by the doctor in the previous segment becomes apparent.

Based on the semantic content of the patient's speech utterances in segment three, we conclude that there also occurs a shift in the **beliefs** of the patient. Until the end of the previous segment the patient resolutely held on to his old beliefs about the situation, following a ***denial*** strategy of coping (see chapter [3]). We argue that the information continuously expressed by the doctor caused the patient to stop holding on exclusively to his old beliefs and to also form new beliefs. This caused the patient to possess two contradicting sets of beliefs about his situation. As a consequence the patient is not sure what to believe, but at this point he suspects that his situation is not as positive as he previously thought. Particularly, the utterance "I though I was cured, but now it sounds like I'm not" indicates that he realizes that his old beliefs do not hold anymore.

Based on the emotion displays of Anger, we argue that the patient has adopted the ***anger*** strategy of coping in this segment. Alternatively, the coping strategy of ***venting emotions*** can be ascribed to the patient. This choice between strategies depends on which categorization of coping strategy one follows. In this segment we see the patient starting to doubt his own beliefs about his situation. This doubt might explain why the patient's coping strategy shifted from ***denial*** to ***anger.***

Segment 4:

The dialogue turns in the fourth segment focus on the coping strategy of the patient now that he is aware of his situation and the closure of the conversation. As mentioned earlier, the fourth segment of dialogue turns partially overlaps with the dialogue turns in the third segment. The utterances after the pause in dialogue turn P7 are included in segment four, while those before the pause belong to segment three.

For continuity and clarity, dialogue turn P7 is completely presented in the transcript of segment three as well as in the transcript of segment four.

P7: *Well I didn't! I don't believe it. (Pause) They have shown you the wrong chart! I'm going to be fine (camera shift). Just need a little medicine and I'll be right as rain.*
D8: *Okay, John. I'll be back later.*

In the speech utterances after the pause in dialogue turn P7, we observe a shift in the emotions of the patient with respect to the emotions expressed before the pause. The displays of Sadness expressed during the first part of the dialogue turn are replaced with displays of *optimism*. These expressions of a positive view of the current situation or of possible situations in the future are somewhat difficult to label and categorize. The most apt way of describing the displays of optimism seem to be as expressions of the emotion **Hope** (about possible future situations) or of an affect disposition, namely *optimistic*. Consequently, we labeled the patient's prosody during the last three utterance in P7 as displays of both, i.e. Hope as taken from the second category and optimistic, which falls into the third category.

The patient's facial expression during the last three utterances is quite neutral in appearance, i.e. it does not contain any displays of emotion, but it confers an approach of *calmness* and *confidence* (at least an attempt to) towards the doctor. Facial expressions of affect dispositions are not included in the annotation scheme and subsequently these displays are not labeled or categorized. We only included the description of the facial expression as a side note which might be used in the categorization of other displays. After the camera shift, we can observe the doctor, whose facial expression indicates he is feeling Sorry-for the patient and tries to *express his sympathy* to the patient.

Just after the pause the patient turns his head away from the doctor, as if he is trying to distance himself from the doctor and from the conversation. This action can be seen as an indication of a change in the interpersonal relation between the patient and the doctor, a decrease of **liking**. The description of the head turn can be placed in category five.

All three of the patient's utterances performed after the pause are of the `behavior type` *`statement`*. The shared `intended effect` of these three utterances is to *convince* the doctor, and perhaps the patient himself, that the situation is not bad at all. We labeled the utterances as having *inform* communicative functions but added the following note to the descriptions: The patient does not only intend for the doctor to know (i.e. form the belief) that the patient believes he is going to be alright, but for the doctor to also believe the patient is going to be alright. The intended effects of the utterances correspond with the coping strategy of ***denial***, displayed earlier in the conversation by the patient. In addition, we interpret the patient's head turn as an indication that the patient does not want (i.e. has the **intention**) to give the doctor the opportunity to correct him. Furthermore, it also indicates that the patient does not want to continue the conversation. The intended effect of the doctor's expression of the emotion Sorry-for is to *inform* the patient about his (i.e. the doctor's) emotional feeling in order to comfort the patient.

The prosody of the utterances in dialogue turn D8 suggests that the doctor reluctantly accepts the decision of the patient to end the conversation for now. We labeled this expression as a *reluctance* affect disposition and place the behavior in the third category. The doctor's face continues to express displays of Sorry-for, started in the previous dialogue turn. Dialogue turn D8 also contains the most expressive gesture in the video clip. During the dialogue turn, we observe the doctor placing his hand on the shoulder of the patient. Based on the context of the bad news situation and the doctor's displays of Sympathy (i.e. a token of the Sorry-for emotion type) in the previous dialogue turn (P7), we also interpret this behavior as the doctor *expressing his sympathy* towards the patient in an attempt to comfort him. The behavior types of the two utterances in this dialogue turn are labeled as respectively a `statement` and an *expressive*. The expressive indicates a goodbye on the part of the doctor.

The communicative function of the first utterance can be interpreted as an *agreement* of the previous *statements* of the patient. While the doctor obviously does not agree with the patient's perception of the situation, this utterance is used by the doctor to maintain a good **rapport** between him and the patient. The second utterance in dialogue turn D8 has the intended effect of *closing the conversation*. Furthermore, based on the semantic content of the speech utterance it becomes clear that the doctor means to continue this conversation at a later time.

The expression of Anger contained in the conversational behaviors in the previous segment have been replaces in this segment by expressions of Sadness. This might indicate that the patient has become aware of the bad news situation he is in. However, analysis of the semantic content of the utterances performed in the conversational behaviors in this segment indicate that the patient has re-adopted the coping strategy of ***denial***, even more strongly than before. The utterance "I don't believe it" in dialogue turn P7 indicates this quite clearly.

The doctor's response to the patient's ***denial*** is to leave him be, something that is not suggested in a protocol such as S.P.I.K.E.S.. Protocols for delivering bad news indicate that it is essential that the bad news is made clear to the patient. The doctor needs to aim for an ***acknowledgment*** by the patient of his understanding of the bad news.

5.2.3 Discussion

The annotation of the conversational behaviors in the video clip has produced a large number of labeled descriptions of the expressions and displays the interlocutors showed during the conversation. In particular, the conversation contained a considerable number of emotional displays, as was to be expected given the topic of the conversation. In addition to the displays of emotions, the number of intended effect descriptions was high as we aimed to label the communicative functions of each conversational behavior.

Based on the observation and the annotations of the displays in the conversational behavior performed in the video clip, we make assumptions about which internal state features the interlocutors experienced during the conversation. As we aim to model the cognitive processes that process and generate conversational behaviors, we are

particularly interested in the **beliefs**, **intentions** and **emotions** of both the speaker and the listener and in the interlocutors' thoughts about the **social relation** between them.

Because the video clip is a simulation of a bad news conversation, the internal state features expressed in the conversational behaviors are obviously not actually experienced by the actors displaying them. Although this is the case, the fact that the video is used as a training tool and contains professional actors reassures us that the enactment is a sufficiently realistic portrayal of an actual bad news conversation. As such, we treat the displays as though they are expressions of genuinely experienced internal state features.

In addition, we presuppose that the displays in the conversational behavior are the actual expressions of internal state features related to the situation and that no form of misdirecting behavior is being performed (e.g. lying, bluffing) during the conversation. If this premise holds, it is quite apparent that the displays of emotions relate to the various elements of the emotion state, i.e. the interlocutor's emotions. For example, an expression labeled as a *display of Anger* is related to the speaker's cognitive experience of **Anger**, which makes up part of the speaker's internal state. The same holds for *displays of Sadness* relating to internal experiences of **Sadness**, *displays of Surprise* relating to internal experiences of **Surprise**, etc.

For other types of internal state features, such as **intentions**, it is more difficult to describe how the expressions in the conversational behaviors relate to them. The labeled `communicative functions` of the conversational behaviors allow us to make assumptions about the intentions they express, but unlike the elements of the emotion state, the set of intentions in an internal state is much less well-defined. Consequently, the assumption about the intention underlying a `communicative function` is derived from the label given to the conversational behavior together with the semantic content of the utterance. For example, the communicative function of the utterance "The surgeon removed the tumor." is to *inform the listener*. The `semantic content` of the utterance provides the information about what the speaker wants (i.e. has the intention) to 'inform the listener' about. In the example, this information is that the surgeon removed the tumor. The intention underlying the conversational behavior can thus be described as: **'Intention: *inform the listener about the fact that the surgeon has removed the tumor.*'**

Beliefs that underlie conversational behaviors are often also difficult to determine. In order to make assumptions about which beliefs hold in the internal state during the performance of the conversational behavior, we use information from several sources. Some of the `communicative functions` operate under the premise that certain **beliefs** must hold in the internal state of the speaker in order for the speaker to perform the behavior. This becomes clear from the definitions found in the DIT++ taxonomy of communicative functions (Bunt, 2009). For example, the definition of the `communicative function` *agree* is: "The speaker believes that the listener believes the semantic content to be true". The semantic content mentioned here is the semantic content of the utterance performed previously by the listener. E.g. the listener said "I think (i.e. **belief**) I'll be better soon." The speaker responds by saying "yeah, you'll be alright." indicating he believes that the listener believes the first utterance and

thus agrees with him. Note that in addition to the agreement, the speaker's utterance can also have the `intended effect` of *informing* the listener about the fact that the speaker also believes himself that the listener is going to be alright.

Ergo, assumptions about the beliefs underlying the conversational behavior can be based on the communicative functions. The exception to this is when beliefs are explicitly stated in an utterance, in which case the labeled expression, together with semantic content are sufficient to describe the beliefs. For example, this holds for the utterance "I think I'll be better soon."

In order to make assumptions about the internal state features that represent the **social relation** between the interlocutors, again several sources are used. On the one hand, the semantic content of an utterance provides clues on the social relation. The choice of words is a good indicator of the speaker's **liking** of the listener, while the *degree of politeness* might indicate the social roles of the interlocutors. On the other hand, the way how the conversational behaviors interact with each other and with regard to the entire conversation can also indicate how the interlocutors perceive each other socially. If there exists a good **rapport** between the interlocutors, the dialogue acts they perform will be more likely to complement each other. We observed that each conversational behavior in the video-clip is an appropriate response to the conversational behavior preceding it. Most of the related conversational behaviors fit together as part of an adjacency pair, such as a *greeting* followed by *return greeting* and a *question* followed by an *answer*. Determining whether or not conversational behaviors belong together, is primarily done based on the `behavior type` of the conversational behaviors and their `semantic content`.

Furthermore, observation of the conversational behaviors in the video clip reveals that the dialogue follows a normal conversational structure, i.e. it contains an opening, a body and a closure. In addition, the turn management of the dialogue turns suggests a high level of *politeness* and *attention* (e.g. there are interruptions or long pauses between dialogue turns). However, these phenomena were not studied in further detail.

Although the annotation of the video clip assists us in making assumptions about the internal state features underlying bad news conversations, this study also has several limitations. One of them is that the amount of data analyzed is too small to be able to make generalized assumptions about internal states underlying behaviors in bad news conversations. As we only annotated and analyzed the conversational behaviors of one bad news conversation, the findings and assumptions only cover the individual conversational behaviors performed in the video-clip, nothing beyond that. After the annotation it became apparent that the set of `communicative functions` needs to be extended. More specifically, the intended effects of expressions of emotions cannot be properly described. Especially displays of emotions that are aimed at causing a change in the internal state of the listener lack a proper description / label.

Furthermore, one might question the validity of the annotations as they have been made by only one annotator. In this sense, the annotations and subsequently the assumptions that are made about the conversational behaviors suffer from similar flaws as Kübler-Ross' model (see section [3.2.3]). That is, the annotations of the displays and expressions can not be considered to be sufficiently objective to give

credit to the assumptions based on them. To validate the annotations of this particular bad news conversation, it needs to be annotated by more people.

5.3 Questionnaire

In addition to our own analysis of the conversational behavior performed by the interlocutors in the bad news video-clip, a second analysis of the conversational behaviors was done by means of an online questionnaire. This second analysis was done with two specific purposes in mind. The first purpose was to gain insight in what kind of thoughts and feelings the interlocutors conducting the bad news conversation were experiencing, as assumed by people who are not specifically educated in describing internal state features. We were interested in how participants would intuitively relate the displays and expressions visible in the conversational behaviors, to features of the internal states of the interlocutors and how the participants would describe those features.

The second purpose of the questionnaire was to see if the assumptions we made regarding the internal state features underlying the displays and expressions, were consistent with those of the participants. If so, the validity of the assumptions could be confirmed.

5.3.1 Setup of the questionnaire

The online questionnaire was set up in the following manner. The video-clip of the bad news conversation was cut into small video-fragments. Instead of dividing the conversation into four segments as we did for the annotation, each video-fragment consists of only a single dialogue turn in the bad news conversation. The dialogue turn could be performed by either the doctor or by the patient. Interruptions of the speech utterances, caused by things such as backchannels or acknowledgments, were not regarded as turn switches. The cutting of the video-clip yielded a total of sixteen video-fragments that then were used in the questionnaire. In total fourteen participants completed the questionnaire. Five of them were female and nine of them were male. The average age of the participants was 31 years and all participants possessed an academic level of education. The participants were asked to watch each video-fragment and closely observe the conversational behaviors performed by the two interlocutors in the fragment. The video-fragments were placed in a chronological order and could be viewed multiple times if desired. After viewing a video-fragment, the participants were asked to answer one to three open questions related to the observed conversational behavior before continuing to the next video-fragment.

If we disregard whose conversational behavior the questions were being asked about (i.e. the doctor or the patient), the total number of different questions used in the questionnaire is five. We expected that these questions would be sufficiently extensive to provide us with the relations between the interlocutors' internal states and the conversational behaviors performed in each specific video-fragment. As a result, similar questions were posed regarding the behaviors in different video-fragments. Some examples of questions that were used in the questionnaire are:

- What do you think the doctor wants to achieve with his behavior?
- Please describe what you think the patient is thinking and feeling at this moment?

In total, the questionnaire consisted of 36 questions relating to conversational behaviors performed in the sixteen video-fragments. See appendix [C] for the entire list of questions used in the questionnaire.

We used open questions in order to avoid possible bias or susceptibility towards terms used to describe features of the internal state. A possible alternative to open questions would be to provide the participants with the labels and descriptions from the annotation scheme used in the first analysis of the conversation. However, in that case the questionnaire would not provide us with the descriptions laymen would use to describe the internal state features.

The participants could answer the questions either in Dutch or in English and they were encouraged to use their own words to describe their observations and assumptions. The questions could be related to the conversational behavior performed by either the doctor or the patient, regardless of whose dialogue turn it was. The reason for also asking questions about the internal state of the listener was because we were also interested in finding out what the participants thought about the effect the speech utterance of the speaker would have on the listener.

5.3.2 Findings of the questionnaire

In this section we present a selection of the answers provided by the participants of the questionnaire. The most similar answers to each question are mentioned, as are the answers that deviate strongly from the expected answers (i.e. answers that correspond with those descibed in section [5.2]). In order to provide a clear image of the context of the questions with regard to the conversational behaviors, the transcripts of the speech utterances performed in each dialogue turn and the related questions are presented.

Doctor1 (D1): *I'm about to visit John Filpot. He is recovering from surgery. He had a bowel tumor removed and evidence of metastases. This means that John's cancer is incurable. Now, I know the surgeon has talked with John, but I don't know if John really understands his condition.*
Q1: Please describe what you think the doctor is thinking and feeling at this moment.

Although Q1 inquires after the participants' thoughts about the internal state features of the doctor, some of the participants reported about the doctor's internal state features by describing the displays they observed. This demonstrates nicely how the participants formed their thoughts about the interlocutors' internal states, namely by associating observed displays with internal state features. As shown in the previous section, we adopted a similar approach to make assumptions about the internal state features based on the annotations.

In the answers to Q1, the majority of the participants reported two emotions they thought the doctor might be experiencing, namely **Sadness** and **Concern for** the pa-

tient. Furthermore, the majority of the participants also reported they thought that the doctor was wondering (i.e. thinking) about the patient's knowledge about the situation and that he is reluctant to talk to the patient. These reports are in line with the assumptions that were made based on the findings of the annotation.

Patient1 (P1): *I'm feeling pretty good. Yeah I got some pain, but the surgeon got my tumor. It was attached to my bowel. But the most important thing is that the tumor is gone. I'll probably be out of here in a few days, once they can remove these tubes and get me on some decent food. I'm sure everything is going to be okay. No biggy, as my son says.*
Q2: Please describe what you think the patient is thinking and feeling at this moment.

The answers to Q2 show a high consensus between the participants. The strongest consensus was that the patient is experiencing the emotion **Relief**. In addition, the majority of the participants reported they thought the patient comes across as *positive* and that he **believed** the state of his situation was good.

D2: *Hey John.*
Q3: Please describe what you think the doctor is thinking and feeling at this moment.
Q4: What do you think the doctor wants to achieve with his behavior?
Q5: What do you anticipate to be the effect of the doctor's behavior on the patient?

If we look at the answers to the questions related to dialogue turn D2, we see that most participants agree in thinking that the doctor is feeling *uneasy / reluctant* to talk to the patient. Also, there is a lot of agreement about what the participants think about what the doctor wants (i.e. has the intention) to achieve with his conversational behavior. The two **intentions** that are ascribed most often to the doctor by the participants are that the doctor wants to 'initiate a serious conversation' and that the doctor wants to 'set the patient at ease' so that he will engage in the conversation.

As described in chapter [3], the first step in the S.P.I.K.E.S. protocol indicates that the bringer of bad news (i.e. the doctor) needs to **Set up** the conversation. The doctor can achieve this by preparing himself mentally (see section [3.2.2]) and by trying to increase the **rapport** between himself and the patient. The first ascribed intention expresses that the doctor believes the conversation is going to be serious, while the second intention indicates that he wants to strengthen the social relation with the patient. Both ascribed intentions are thus in line with the first step in the S.P.I.K.E.S. protocol.

Based on the answers, we deduce that most participants expected the behavior of the doctor to affect the emotions of the patient, causing him to experience Surprise and Fear. The argumentation given by most participants for these answers is that the doctor is displaying a Sad and Worried facial expression and tone of voice.

P2: *Hello doctor Rupert.*
Q6: What do you think the patient wants to achieve with his behavior?

We only posed a single question concerning the conversational behavior in dia-

logue turn P2 because we believed that posing any of the other questions would result in the same answers that were given in response to Q2. The answers to Q6 show that the majority of the participants think the patient wants to (i.e. has the **intention** to) 'engage in the conversation with the doctor and give him his full attention'.

Some of the participants reported that this **intention** of engagement and allocation of attention are expressions of an increase in the aspects regarding the social relation between the patient and the doctor. This is apparent in the answer: "(The patient) indicates he's going to listen. Showing he is polite to someone who has cured him."

D3: *I'm sorry to hear the news. We were hoping that surgery and a course of chemo would cure this tumor.*
Q7: What do you think the doctor wants to achieve with his behavior?
Q8: What do you anticipate to be the effect of the doctor's behavior on the patient?

The participants' answers to Q7 show a lot of variety in the **intentions** ascribed to the doctor underlying the behavior in dialogue turn D3. This is a good example that a single conversational behavior can express more than one intention. The various intentions ascribed to the doctor focus on *influencing different features in the patient's internal state*.

Several participants state that the doctor's intention is to *inform* the patient about the beliefs of the doctor (which is knowledge about the state of the patient's situation) in an attempt to change to patient's beliefs. According to other participants, the doctor's intention is to try to comfort the patient (i.e. *Alter the patient's emotion state*) and to *express the doctor's Sympathy*.

All of the answers to Q8 are roughly similar in content. The participants anticipate that the doctor's behavior will cause the patient to experience Surprise or Confusion. Some participants indicate the patient might also experience Fear or Sadness as a result of the doctor's conversational behavior. It seems likely that this is the result of the participants knowing the internal state of the patient beforehand (namely that he thinks that everything is okay), and that the expected emotion states are caused by the conflict between the patient's old beliefs and the information the doctor is providing.

The assumptions made by the participants about the patient's intentions are comparable to the assumptions made based on the annotations of this conversational behavior.

P3: *What do you mean?*
Q9: What do you think the patient wants to achieve with his behavior?
Q10: What effect do you think the doctor's previous behavior had on the patient's behavior in this fragment?
Q11: Please describe what you think the patient is thinking and feeling at this moment.

From the answers to Q9, we find that the participants are quite like-minded in their beliefs about what the patient's **intention** is that underlies the patient's conversational behavior in dialogue turn P3. According to the majority of the participants,

the patient's intention is to *get more information* from the doctor to resolve his Confusion about the situation.

Question Q10 was used to see if the participants could indicate the relation between the conversational behavior performed in dialogue turn D3, the internal state features of the patient and the conversational response performed in dialogue turn P3. The majority of the participants stated, in answer to Q10, that they thought the doctor's previous conversational behavior caused Confusion and Surprise in the patient. Only a couple of the participants also indicated that the state of the patient's emotions caused the patient to perform this particular conversational behavior. An example of such an answer is: "it got the patient confused, causing him to ask "what do you mean?"".

Most of the participants answer Q11 by stating they think the patient believes something is wrong or that he is missing some information about his situation. In addition, they report that they think the patient is feeling a lot of Confusion and Surprise.

D4: *When a tumor has spread through the abdomen, like yours has, there is a limited amount we can do for you.*
Q12: What do you think the doctor wants to achieve with his behavior?
Q13: What do you anticipate to be the effect of the doctor's behavior on the patient?

A large number of the participants answer question Q12 by saying they think the doctor has the intention of explaining the state of the patient's situation by *providing more information*. Some of the participants extend their answer by adding that they think that the doctor is *providing more information* to clarify the patient's Confusion. This ascribed intention corresponds to step four of the S.P.I.K.E.S. protocol, which is **giving Knowledge and information** to the patient to attempt to make the bad news situation clear.

The conversational behavior in dialogue turn D4 is anticipated by a large number of the participants to cause the patient to experience more confusion and surprise. Concerning how the patient will deal with the conversational behavior of the doctor, there is no clear consensus between the participants. One participant expected the patient would start crying (which we interpret as an expression of the ***depression*** strategy coping), while other participants anticipated the patient would still be confused and as such will not accept what the doctor was telling him. It seems that at this point there are not clear cues yet to determine how the patient will deal with his situation.

P4: *What do you mean, spread through the abdomen like mine has? The surgeon said he had removed the tumor.*
Q14: What do you think the patient wants to achieve with his behavior?
Q15: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q16: Please describe what you think the patient is thinking and feeling at this moment.

If we look at the answers provided to the questions dealing with dialogue turn

P4 we see that the participants ascribe the following internal state features to the patient in answer to Q16. The patient is experiencing feelings of conflict between his old beliefs (i.e. 'the situation is fine') and the new beliefs he has formed based on his thoughts about the beliefs of the doctor (i.e. 'the situation is not fine'). These conflicting sets of beliefs in turn lead to the formation of even newer beliefs: e.g. 'There has been made a mistake by the doctor / What is the doctor talking about, the surgeon said I was okay / This is not correct.'

The answers to Q14 show that the majority of the participants believe that the patient has one of two intentions that are associated with dialogue turn. The first `intention` ascribed to the patient is to *`get more information`* about his situation from the doctor. The other `intention` is to *`provide information`* to the doctor in order to clarify the confusion. The emotions of the patient are slowly turning from **Confusion** and **Surprise** to **Worry / Fear** and to some extend **Anger**. This might be an indication that the patient is getting a more clear idea about the situation.

With respect to the `expected effect` on the doctor, the participants report that they anticipate that the doctor will form the intention to *`provide extra information / explain the situation`* and experience **Surprise, Confusion** and **Embarrassment**.

D5: *Yes, The primary one. But he couldn't get all the places it had invaded, we need to give you chemo for that.*
Q17: What do you think the doctor wants to achieve with his behavior?
Q18: How do you think the doctor's behavior will have an effect on the patient?
Q19: Please describe what you think the doctor is thinking and feeling at this moment.

The intention that is ascribed to the doctor by a large number of the participants is that he wants to clarify the situation to the patient by *`explaining`* and *`elaborating`* his previous behaviors. It is likely that this interpretation of the doctor's behavior is caused by the patient's previous behaviors and by the new beliefs the doctor has formed about the beliefs of the patient (i.e. 'the patient does not know the situation is bad') based on these behaviors. The doctor's emotions that are influenced by his new beliefs are those related to **Confusion**, **Sorry for** the patient (i.e. **Empathy**) and **Embarrassment**.

All the participants' answers to Q18 show that it is expected that the patient's emotions will strongly be affected by the conversational behavior of the doctor. However, the specific emotions that are believed to be affected vary.

Participants mention the emotions of **Fear, Confusion, Sadness, Anger** and **Disbelief** to be part of what the doctor is thinking and feeling while performing the conversational behavior in D5.

P5: *What do you mean invaded? You keep talking about this chemo-stuff.* ***Q20:*** *What do you think the patient wants to achieve with his behavior?*
Q21: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q22: Please describe what you think the patient is thinking and feeling at this moment.

The answer most often given in reply to Q20 is that the participants think the patient intends to *get more information and clarification* about the situation from the doctor. As a consequence to this conversational behavior from the patient, the participants expect the doctor will oblige the patient by *providing more information* about the situation. Also the doctor is expected to maintain his emotion of **Surprise**.

All participants' answers to Q22 were descriptions of the patient's emotions. Various emotions were ascribed to the patient, with **Worry / Fear** being the most prominent, followed by **Anger**. Other emotions that were mentioned in the answers are **Surprise, Confusion** and **Sadness**. It is possible that the participants were primed to only ascribe emotions to the patient instead of also other internal state features, as they described the expectations of emotions in the answers to Q18.

D6: *Didn't the surgeon talk with you?*
Q23: What do you think the doctor wants to achieve with his behavior?
Q24: What do you anticipate to be the effect of the doctor's behavior on the patient?
Q25: Please describe what you think the doctor is thinking and feeling at this moment.

The participants reported that the assumed underlying reason of the doctor's conversational behavior is to find out what the surgeon exactly has told the patient. In other words, the `intention` that was ascribed most often to the doctor in this dialogue turn by the participants is to *gather information* about the patient's knowledge of his situation.

Some of the participants also mentioned other intentions in their responses to the question about the patient's internal state. These answers describe the doctor's intention to *(dis)confirm* his **beliefs** about the patient's situation.

Furthermore, a large number of the participants responded that they anticipated that the patient's response to this conversational behavior of the doctor would be (i.e. have the behavior type) an *affirmation*, i.e. provide an answer to the doctor's question. Additionally, the participants anticipate that the patient will become **Irritated / Angry** at the doctor, but also **Relieved** that the doctor finally realizes the source of the misunderstanding. One participant reported that he expected the patient to doubt the things the surgeon has told him and inquire about them with the doctor.

The internal state features (i.e. what he is thinking and feeling) that are ascribed to the doctor are **Confusion, Anger at the surgeon** and also **Irritation** at the patient for not understanding what he is saying. The participants also report that they think that the certainty of the doctor's beliefs about what the surgeon has told the patient decrease rapidly.

P6: *Yeah but he didn't say anything about chemo or invasive stuff. What the hell is going on here? I thought I was cured, but now it sounds like I'm not.*
Q26: What do you think the patient wants to achieve with his behavior?
Q27: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q28: Please describe what you think the patient is thinking and feeling at this moment.

The participants report that they think the primary intentions of the patient are to *express his emotions*, foremost displaying his **Anger** and **Frustration**. From this we conclude that these emotions have the highest intensity values at this moment. According to the participants, the emotions **Unhappiness**, **Surprise**, **Confusion** and **Fear** are believed to still be present in the patient (i.e. they are the lingering result of earlier conversational behaviors), but in less degree. Consequently, we argue that the intensity values of these lingering emotions are less than those associated with the emotions **Anger** and **Frustration**.

Some of the beliefs ascribed to the patient based on the conversational behavior in P6 indicate that contrary to what he believed before, he believes his situation is not okay. The feelings that are ascribed to the patient correspond with the described intentions. That is, **Anger** and **Frustration** are most often / strongly perceived and subsequently described by the participants followed by **Unhappiness**, **Surprise**, **Confusion** and **Fear**.

From the answers to Q26 we take that the participants think that besides the intention to express his feelings, the patient also has the intention to find out what the status of his situation is and what the doctor's beliefs are. We argue that the participants may have formed these thoughts partially because the conversational behavior of the patient consists for a large amount of questions.

The doctor is expected by the participants to feel **Embarrassed** about the confusion and worried about the patient. In addition, the doctor is believed to have the `intention` to *explain the situation* (about the patient's health) more clearly and in more detail (as he believes the patient is now more aware of his situation).

D7: *I'm sorry John. I thought you knew.*
Q29: What do you think the doctor wants to achieve with his behavior?
Q30: What do you anticipate to be the effect of the doctor's behavior on the patient?
Q31: Please describe what you think the doctor is thinking and feeling at this moment.

Seemingly it is quite difficult to pinpoint what the doctor wants to achieve with this behavior. The participants describe a variety of intentions underlying the conversational behavior of the doctor. Some of the intentions included in the participants' answers are that the doctor wants the patient to maintain **Hope** (i.e. *alter the patient's emotion state*), the doctor wants to *express his Concern* for the patient, the doctor wants to *express his Sympathy* to the patient and the doctor wants to *explain* to the patient that he (i.e. the doctor) believed the patient was already aware of his situation. Although each of the participants' responses only ascribed a single intention to the doctor, we argue that it is possible and plausible that more than one of the mentioned intentions is expressed in the same conversational behavior.

The participants' expectations about which the effect the doctor's conversational behavior has on the patient include the following **beliefs**: The majority of participants predicts that the patient will become angry at the doctor and that his response behavior will contain various displays of **Anger**. From these expectations we draw the conclusion that the doctor's conversational behavior will cause the intensity value of the patient's **Anger** emotion type to increase and that the `communicative function`

of the patient's response behavior is to *express this emotion*.

Another expectation formed by some of the participants regarding the effect of the doctor's behavior is that the patient's response behavior will express his coping strategy of ***denial*** (i.e. through his behavior, the patient will reject the information provided by the doctor).

According to the answers given in response to Q31, the participants indicate that they primarily think the doctor is experiencing emotions at this point. Feelings of **Embarrassment, Sympathy** and **Confusion** are most prominently ascribed here.

P7: *Well I didn't! I don't believe it. (Pause) They have shown you the wrong chart! I'm going to be fine* (camera shift). *Just need a little medicine and I'll be right as rain.*
Q32: What do you think the patient wants to achieve with his behavior?
Q33: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q34: Please describe what you think the patient is thinking and feeling at this moment.

Analysis of the answers provided in response to question 32 indicates that most of the participants think the patient want to contradict the information the doctor has provided in his earlier behaviors. We argue that the intended effect of such contradicting behavior may be to *alter features in the recipient's internal state* (i.e. the doctor's **beliefs**) or *the speaker's internal state* (the patient's own **beliefs**) or *both their internal states*. Based on the answers of the participants and the information presented in chapter [3], we conclude that the patient is following the ***denial*** coping strategy. The participants' answers describe a distinction between two types of ***denial***: The first type is where the patient directly states that the doctor is wrong and where he is maintaining the belief he is going to get better. The second type is more of a 'soft' denial, where the patient acknowledges the situation, but tries to mitigate or downplay the severity of the situation.

Another function the participants ascribe to the patient's conversational behavior is the expression of his intention to *close the conversation*. According to the participants, the patient tries to achieve this intended effect by ignoring the doctor (e.g. turn away, break eye contact).

Most participants anticipate that the patient's conversational behavior will affect the doctor in one of two ways. They expect that the doctor will either continue to try and explain the situation to the patient or leave the patient alone to allow him time to process the new situation.

Other expected effects of the patient's conversational behavior are adjustments to the doctor's emotions. More specifically, it is anticipated that there is an increase in his worrying about the patient, his feeling of **Discomfort** and his feeling of **Sadness**.

In answers to question 34, the participants indicate that they think the patient's emotions shift focus again. The new emotions he is believed to be experiencing are **Sadness, Disappointment, Helplessness, Disbelief** and **Anger**, with some residual effects of **Surprise** and **Confusion**.

D8: *Okay, John. I'll be back later.*
Q35: What do you think the doctor wants to achieve with his behavior?

Q36: Please describe what you think the doctor is thinking and feeling at this moment.

The answers given to question 35 show that the majority of the participants believe that the `intended effect` of the doctor's behavior is to *`close the conversation`*. They substantiate their belief about this effect by arguing that the doctor is aiming to close the conversation with the intention to allow the patient time to assess his situation or possibly coming to terms with it. We argue that these statements indicate that the `intended effect` of a conversational behavior does not have to be the same as the underlying **intention** (i.e. a feature of the internal state) but that they are inextricably related to each other.

In addition to this intended effect, other intended effects associated with the doctor's conversational behavior are that the doctor is trying to *`comfort the patient`* and *`expressing his support`*. These ascribed intended effects are substantiated by the participants' assumptions that the doctor is experiencing feelings of **Sadness** and **Sympathy** with the patient. The participants express this by mentioning they observed a display of these emotions (i.e. the doctor by putting his hand on the patient's shoulder). The display and the associated emotions are ascribed to the doctor in the answers to question 36.

5.4 Conclusions

In the introduction of this chapter, we argue that by observing and interpreting conversational behaviors in a bad news conversation we can make assumptions about which internal state features of a speaker hold at the moment he is performing a conversational behavior in such a conversation. In addition, we argue that we can make assumptions about how that conversational behavior might affect the internal state features of the listener and subsequently influence his response behavior.

In order to make founded assumptions, a thorough examination of the various properties of conversational behaviors was deemed necessary. Presented in this chapter are the two analyses that have been performed on the conversational behaviors that were expressed in the bad news conversation shown in the Pallium video-clip. The purpose of these analyses was to find out how the various features of the internal state of both interlocutors, and the processes associated with them, are related to the conversational behaviors performed in the bad news conversation shown in the video-clip.

The first analysis of the conversational behaviors was done based on an annotation we made concerning the various kinds of displays that express one or more features of the internal state. In addition, the dialogue act type (i.e. the communicative function or the intended effect) and behavior type of each individual conversational behavior were annotated.

The set of internal state features, the `dialogue act types` and `behavior types` that could possibly be annotated was constructed based on the information we gathered from various studies and which is presented in chapters [2] and [4]. The collection of internal state features consists of **beliefs, desires, intentions** and the **26 emotion types**. A complete list of the dialogue act types and behavior types, which

could be annotated, is presented in section [5.2.1].

In addition to specified annotation labels, we also included pointers in the annotation scheme which annotators might use to describe the perceived displays. This was done because we anticipated that a sizable portion of the possible displays could not be properly explained by the provided annotation labels. The alternative was to specify the annotation labels in such a way they would provide a more general description of the display. However we decided against this option as we believe that by generalizing the annotation labels we would lose valuable information regarding the relation between displays and the underlying internal state features.

From the first analysis, we conclude that emotion types can be associated most easily and most directly to conversational behaviors. We argue that this is because emotion types are expressed more explicitly through displays in the conversational behaviors than any of the other internal state features. In addition, the intentions that underlie a conversation behavior and the beliefs that are associated with that intention are derived more often from interpreting the dialogue act type, the behavior type and the semantic content of the conversational behavior than from interpreting explicit displays.

However, as we only annotated and analyzed the conversational behaviors of one bad news conversation it is the question whether the selected annotation labels were sufficient to describe the displays and their associated internal state features. Therefore, a second analysis was performed.

The basis for the second analysis consisted of an online questionnaire. In addition to analyzing the answers to each question, we also look at the similarity of the answers and the agreement between the participants. If a question received many similar answers, we assumed that those answers described the most likely and most accurate description of the speaker's or the listener's internal state. Note that we took into account the possibility that each answer described only part of the internal state asked. In other words, while we assume that the internal state features ascribed to the doctor or the patient by the majority of the participants of the questionnaire (as the describe in their answers) are accurate (i.e. that those features indeed hold in the doctor's or the patient's internal state), we also allow the possibility that internal state of the doctor or the patient can contain additional features.

Analysis of the answers to the questionnaire has led us to the following conclusions. The interlocutors already possess a large collection of internal state features before the start of the conversation. Analysis of the answers to questions Q1 and Q2 give us some idea about what is included in the background of knowledge (i.e. **beliefs**), **goals** and starting **emotions** to perform a bad news conversation.

The answers provided by the participants of the questionnaire show that during a large of the conversation (dialogue turns P3, D4, P4, D5, and P5) both the patient and the doctor are filled with feelings of **Confusion** and **Surprise**. In dialogue turns D6 and P6 these emotions change as a result of a better understanding of the situation. This indicates quite nicely that the focus of the emotion state may shift from one emotion type to another during a conversation.

From the answers, we also conclude that many of the doctor's and the patient's **intentions** aim to resolve their respective emotions by trying to *gain more information*

and clarification from the interlocutor. This approach of gathering more information is the result of having **intentions** to form new **beliefs** about the situation with a high certainty.

Answers to the questions about dialogue turns D7, P7 and D8 indicate that the **goals** and **intentions** of both interlocutors have shifted. At the end of the conversation the intentions of both the doctor and the patient are to *provide the other with new information*.

The answers to the questionnaire did not provide us with much information about which coping strategies the patient was using. Only the last couple of questions provided answers in which the participants describe that the patient is coping by ***denying*** the doctor's words. As described in chapter [3] the ***denial*** coping strategy is one of the stages of dealing with the bad news. However, we also conclude that although the participants do not name the coping strategy of ***focusing on*** and ***venting of emotions*** (described in section [4.3.4]), they do indicate that both the doctor and the patient perform a lot of *displays of emotion*. This can be interpreted as amounting to the same thing.

The answers indicate more clearly that the doctor did not adhere completely to the protocols that were also described in chapter [3]. From the participants' answers we conclude that the doctor only follows steps one and four of the S.P.I.K.E.S. protocol. While the doctor seems prepared for the conversation, his lack of knowledge about what the patient knows causes the bad news conversation to go awry. As a result, the bad news conversation in the video-clip can be used to show how giving bad news can go wrong.

If the doctor wanted to hold the bad news conversation in the correct manner, he should have inquired about the patient's beliefs about the situation and have taken these beliefs into account when determining which conversational behaviors to perform.

If we look at the overall analysis of the questionnaire, we see that the results showed quite a lot of corresponding assumptions about the features of internal states between the participants. In other words, the participants showed quite a lot of agreement in their answers.

As stated previously, we argue that a high agreement between the participants indicates that the descriptions of the internal state assumed in the answers are most likely accurate. After comparing the matching answers of the questionnaire to the assumptions made based on the results of the annotation, many similarities were found. From this we conclude that our assumptions from the first analysis (i.e. the annotation) are probably accurate.

In addition, we found that many of the descriptions from the second analysis (i.e. the questionnaire) show similarities to the labels we used in the first analysis. This holds for the displays observed in the conversational behaviors as well as for the internal state features that have lead to the performance of those behaviors. Based on this observation and the fact that the construction of the labels was based on studied theories and methods, we conclude that the annotation scheme we used for the first analysis is reasonably accurate.

Furthermore, it indicates that even without having studied cognitive theories, lan-

guage processing methods, etc. or possessing any formal training human beings are able to intuitively interpret the displays and properties of a conversational behavior and quite accurately can make assumptions about which features in the speaker's internal state are associated with those displays and properties.

Part II

Constructing the Dialogue Model

6

Chapter 6: Cognitive dialogue model

In earlier chapters we discussed a number of methods and theories that describe one or more aspects of conversational behaviors performed in natural spoken dialogues. Such aspects include: the content and purpose of a conversational behavior, the form of the utterances used in a conversational behavior (e.g. whether the utterance is a question, or a statement) and the internal state features and cognitive processes of an interlocutor that interpret, process and generate conversational behaviors.

In chapter [2], we discussed speech act theory, dialogue acts and the DIT++ taxonomy of communicative functions, which describes the meaning and the function (i.e. the `intended effects`) of a conversational behavior. The DIT++ taxonomy also provides insight into how dialogue acts relate to internal state features other than **intentions**, such as **beliefs** and **goals**. In addition, chapter [2] contains a general overview of the different methods which can be used to model dialogues.

Theories about the kinds of conversational behavior that people perform or should perform in bad news situations are presented in chapter [3]. Moreover, the difficulties that can occur during the processing of conversational behaviors performed in bad news situations are described in detail. Furthermore, chapter [3] described the various ***coping strategies*** that people may follow when dealing with bad news situations.

Chapter [4] covered several methods and theories that describe different types of internal state features and how such features are possibly processed on a cognitive level. Foremost amongst these theories is the BDI-model of practical reasoning (Bratman, 1987). A small number of the presented methods and theories focuses on the modeling of emotions. In particular, the OCC-model (Ortony et al., 1988) was closely studied and explained. Other methods and theories presented discuss how coping mechanisms operate and how these mechanisms lead to the selection of conversational behaviors. Two of such theories are the Ways of Coping (Folkman and Lazarus, 1985) and the COPE approach (Carver et al., 1989), which were studied in detail.

The two analyses presented in chapter [5] showed how conversational behaviors performed in a bad news conversation are perceived and interpreted by human observers. Based on these analyses we make assumptions about the relations between

displays and expressions visible in the conversational behaviors and features of the internal state of both interlocutors.

The information obtained from studying these theories, mechanisms and real life conversations was used to construct a cognitive dialogue model that represents human-like conversational behavior. By constructing a dialogue model that is able to represent the internal state features and the processes involved in the processing of conversational behaviors performed during bad news conversations, we can provide health care workers with the tools to learn about and practice with holding such difficult situations. In this chapter, the cognitive dialogue model we have constructed is presented.

The goal of this model is to provide a basis, in the form of an architecture, from which a dialogue system can be created that processes and performs conversational behaviors in a human-like manner. By using elements and rules that represent the internal state features and the cognitive processes involved in conversations, any dialogue system that is constructed based on the model should be able to provide intuitive insight to the user about why and how the system performed the conversational behaviors it did. An additional motivation for constructing our own dialogue model instead of adhering to existing models has been that, in our opinion, existing models are too limited in the processing of conversational behaviors or only focused on one specific aspect of dialogue processing. By limiting a model or theory to only a single aspect of dialogue processing, one runs the risk of excluding vital relations between different aspects which might influence the functioning of that aspect.

This chapter is structured in the following way: In section [6.1] we present a categorization we made for the conversational behaviors performed in the Pallium video-clip analyzed in chapter [5]. Categorizing the conversational behaviors is done based on the findings of the annotation and questionnaire presented in the previous chapter. The categorization provides us with insight in what the input and output streams of the dialogue model can consist of and how these are related to the representations of the internal states of the interlocutors.

Section [6.2] gives an overview of the components of our dialogue model and presents the structure that underlies the components. Furthermore, section [6.2] describes what the contents of these components are and how these contents are represented. The contents of the components consist of information about the situation that is represented in the participants' internal states, i.e. the interlocutors' **beliefs, goals, intentions, emotions, social views** etc. but also the *formation* and *updating rules* for processing this information.

In section [6.3] we compare our dialogue model to the models, theories and methods we described in chapter [4]. Several of these models, theories and methods form the basis for the majority of the components in our dialogue model. Therefore, we also describe in section [6.3] how the various models and theories have been modified to fit into our dialogue model and how our components differ from them.

In order to be able to validate our theoretical cognitive dialogue model, it needs to be incorporated into a working dialogue system. To that end, we made a simple implementation of the content of the various components in the dialogue model. This simple implementation is described in Section [6.4].

Section [6.5] contains our conclusions about the categorization of conversational behaviors in bad news conversation we have made and the functioning and structure of our cognitive dialogue model.

6.1 Taxonomy of Conversational Behaviors

In order to create the cognitive dialogue model, all the thoughts, feelings and internal processes used during a bad news conversation in both the speaker and the listener need to be represented. As these internal state features and processes are either formed or affected by the conversational behaviors performed during the conversation, we try to structure the conversational behaviors used in a bad news conversation and represent the related internal state features and processes accordingly. We argue that each conversational behavior can be described by means of a set of `behavior properties` that can assist us in determining which internal state features and processes underlie the selection and manner of performance of that particular conversational behavior. Such information can be used by the listener (or system) to understand why the speaker (or user) chose to select and perform the conversational behavior, making it easier to select an appropriate response behavior.

Describing the properties of conversational behaviors requires close observation and interpretation of the semantic content and the physical characteristics (e.g. the configuration of the face, the prosody of the voice, etc.) of the conversational behaviors. The two analyses performed and presented in the previous chapter are good examples of this. Due to time constraints and the complexity of the process, a detailed description of the process of interpreting semantic content is not included in this thesis. Instead, we trusted that the humans performing the analyses were sufficiently proficient in interpreting the content of the offered conversational behavior. Subsequently, the process of interpreting observed conversational behaviors is also left out of the dialogue model as otherwise the scope would become too large.

Based on the findings of the annotation and to a lesser degree the findings of the questionnaire, we argue that conversational behaviors can be described via the following properties:

a) The `intended effect` of the conversational behavior. This is what the speaker aims to bring about through the performance of the conversational behavior. This property is also known as the `illocutionary point` (Searle, 1969), the `communicative function` (Bunt, 1994) or the `speaker's meaning` (Clark, 1996). Often when conversational behaviors are used in a dialogue, their intended effects are to alter something in the internal state of the listener. As mentioned earlier, we argue that the intended effect of a conversational behavior is related to the **intention** of the speaker.

b) The `expected effect` of the conversational behavior. Normally, the expected effect is similar to the intended effect: a speaker expects that his conversational behavior has the effect that he intends to achieve. For example, if the intended effect is to *`gather information`*, the speaker also expects to get the information. Otherwise performing the conversational behavior would be pointless, e.g.

a person does not ask a question if he intends to get an answer but does not expect he will get one. Although the `intended effect` and the `expected effect` are normally similar, they differ in the internal state features to which they are related. The `intended effect` comes from the speaker's **intentions**, while the `expected effects` are contained in the speaker's **expectations** (i.e. a specific type of belief).

c) A specific `configuration of the internal state features` of the speaker that has led to the selection of the conversational behavior performed. The configuration consists of particular, necessary presuppositions and attitudes that inevitably accompany the illocutionary point. It is not sufficient to have an **intention** to form a conversational behavior with an `intended effect`. The internal state of the speaker also must contain certain other internal state features, such as **beliefs** or a certain disposition, which must hold before the conversational behavior will be selected and performed. Determining what these internal state features are is done in two ways. The first way is through experience. If the listener is familiar with the content of the conversational behavior or if he has performed similar conversational behaviors, the listener knows what kinds of internal state features should underlie the speaker's conversational behavior. The second way is to infer the internal state features from the intended and expected effect properties. Based on the `intended effect` of a conversational behavior, the listener can ascribe **intentions** and **desires** to the speaker and based on the `expected effects` he can do the same for the some of the speaker's **beliefs** (e.g. the speaker's **expectations**).

   For example, in order to select an Express empathy kind of conversational behavior, the speaker must have the following feature values in his internal state: the **belief** that the listener is in a bad situation, the **belief** the listener is experiencing negative emotions, an understanding (in the form of **beliefs**) of the bad situation, and the **desire** to inform the listener of the speaker's understanding (which will become an **intention** when the speaker commits to seeing it fulfilled).

   Together, the `configuration of internal state features` and the `illocutionary point` (i.e. the `intended effect`) form the *illocutionary force*, which is discussed in chapter [2]. Note that this property (the configuration) of the conversational behavior only describes the features of the internal state that necessarily must hold before the behavior will be selected.

d) The `behavior type` of the conversational behavior and the consequent relation it has with respect to the behavior types of other conversational behaviors. These other conversational behaviors are either performed earlier or expected to be performed later on in the dialogue. The behavior type or form of the conversational behavior is discussed in chapter [5]. Examples of relations between two conversational behaviors based on their behavior type are: a *`question`* is often followed by an answer in the form of a *`statement`*. The *`question`* may be posed to get an explanation of the content of a prior conversational behavior.

e) Emotions that can be ascribed to the speaker based on interpretation of the conversational behavior. Although the emotions are not actually a property of the conversational behavior, they are expressed through `emotion displays` contained in the conversational behavior, such as the prosody of the speaker's voice or his facial expression. For example, a conversational behavior performed in a loud voice with short sentences, might lead to the ascription of the emotion **Anger** to the speaker.

   It is important to note that the displays of emotions are not necessarily aimed to achieve the intended effect of the conversational behavior, nor are the emotions necessarily presuppositions needed to select the conversational behavior. It is possible that emotions only influence the way a conversational behavior is performed.

   However, if the displays of a certain emotion are aimed to achieve the intended effect of a conversational behavior, then that emotion is needed before the conversational behavior can be selected. For example, if the speaker wants to *inform* the listener that he is feeling sad, the internal state feature **Sadness** must hold and the performed conversational behavior needs to contain *displays of Sadness*.

f) The `social relation` that exist between the two interlocutors according to the speaker, which is expressed through features such as the degrees of **liking**, **trust** and **rapport** the interlocutors have for each other and the social roles established between them (for example tutor-pupil or doctor-patient). Similar to emotions, the social relation is not an actual property of the conversational behavior, but it can be expressed through displays contained in the conversational behavior, such as hugging or close proximity.

   Alternatively, the social relation can also be expressed through the semantic content of the conversational behavior. An example of this is the *degree of politeness* the speaker uses in his speech utterances or the use of formal or informal words. In this case, an annotation is made in property g) about the state of the social relation as it is perceived.

   Similar to the previous property, expressions of the social state are not necessarily used to bring about the intended effect of the conversational behavior, but they can also influence the form of the conversational behavior.

g) The `semantic content` of the conversational behavior. The semantic content contains the factual information regarding the situation, i.e. the objects, propositions and events, that are referred to in the conversational behavior. The semantic content is vital in identifying what it is that is intended to be affected in the listener's internal state by the conversational behavior. The semantic content is similar to the signal meaning Clark (1996) explained in chapter [2].

The properties of the conversational behaviors can be looked at from different levels. The different levels indicate the degree of detail that is provided by a particular property. This is represented in figure [6.1].

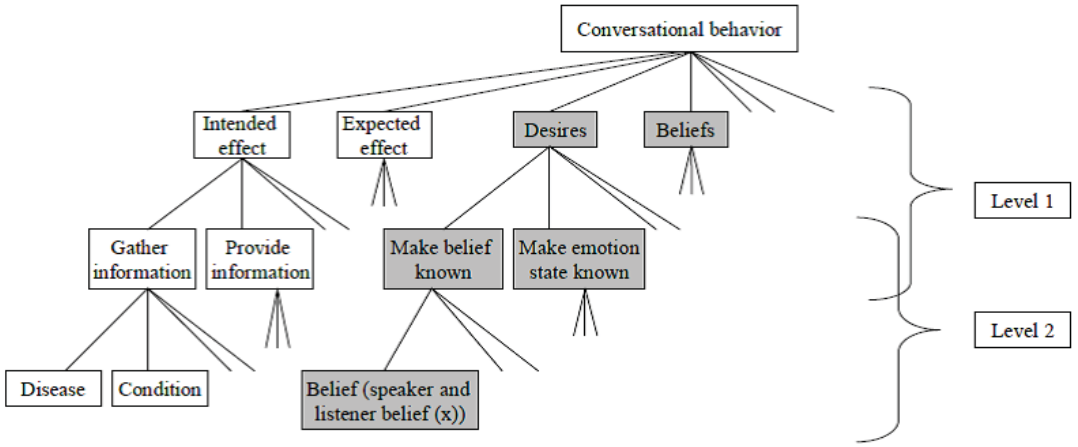

Figure 6.1: Levels of detail of properties and values.

Level 1 describes the `properties of conversational behaviors` and shows the general descriptions of the possible *`atomic values`* for the properties. As mentioned, each property can have various values. For example, the conversational behavior "What is my condition?" has an `intended effect` to *`obtain information`* about a certain situation. The same holds for the conversational behavior "Can this disease be cured?" and for all conversational behaviors that focus on getting information from the interlocutor. As a result we can argue that one possible value that is associated with the intended effect property is *`question`* (i.e. to gather information).

This value of the intended effect property differs from the one associated with a conversational behavior such as "You have a malignant tumor". The atomic value of the intended effect (at level 1) of this conversational behavior is *`inform`*. As can be seen *`question`* and *`inform`* are two different values of the property `intended effect`. A good overview of different kinds of intended effects is the DIT++ taxonomy of communicative functions (Bunt, 2006), which is discussed in chapter [5].

Another example that shows that the properties of a conversational behavior can have different values is the following: the `configuration of the internal state features` that leads to the selection of a conversational behavior varies from behavior to behavior. In figure [6.1], this property of a conversational behavior is represented as the gray area.

The conversational behavior *`show agreement`* (i.e. has that intended effect) is selected based on the speaker's **belief** that he and the listener have a particular belief in common, and the speaker's **desire** to make this known to the listener. These two general internal state features make up the configuration property of this conversational behavior. A different configuration of underlying internal state features is associated with a conversational behavior which aims (i.e. has the intended effect) to express Joy. In this case the configuration of internal state features contains the speaker's **desire** to make his emotions known to the listener and, the **desire** to positively influence the emotions of the listener. Furthermore, the configuration also includes speaker's emotion **'Happiness'**.

Some of the values of the properties described at level 1 can be quite complex. Therefore, we describe some of the level 1 values as property-value pairs themselves. This more detailed representation corresponds to level 2. For example the information that is intended to be *gathered* or *provided* can be different for each conversational behavior that has such an `intended effect`. More specifically, the conversational behavior "What is my condition?" has the intended effect of gathering information about 'the condition of the patient', while the conversational behavior "Can this disease be cured?" inquires after 'the curability of the disease'. The level 1 values of the `intended effect` properties are the same for both conversational behaviors (i.e. *question* / gather information) but the individual (i.e. level 2) values differ. The extended specification of values holds for all the properties of conversational behaviors.

6.1.1 Categorization of conversational behaviors

The wide variety of possible level 2 values makes it difficult to construct a model that is able to process all the different conversational behaviors performed in a dialogue. Therefore we group similar conversational behaviors together, based on the level 1 values of their properties. In order to keep the categorization of conversational behaviors manageable, we limited the conversational behaviors to be grouped to those performed in the Pallium video. The two analyses of the Pallium video-clip that are presented in chapter [5] already revealed many of the properties of the conversational behavior performed in the bad news conversation. In addition, we use the findings and the assumptions we made based on the findings to fill in the undetermined values, both for level 1 and level 2. The determined properties are used to see what the common denominators for the conversational behaviors are and which values should be used to determine the categories.

Because the terms "speaker" and "listener" quickly become confusing when used in the context of a dialogue, we use the variables A and B to indicate the two interlocutors in the description of the dialogue model. We consider A as the person from whose perspective the conversational behaviors in the dialogue are being observed and processed. In other words, A represents the one whose mental model is represented by the dialogue model. Person B is the one with whom A is talking.

The function of conversational behaviors within the dialogue model can be seen from two perspectives. Note that we do not mean the `communicative functions` of the conversational behaviors here, but rather the way the conversational behaviors operate with respect to a person in a conversation. The first perspective regards the conversational behaviors (performed by B) as part of the input of the dialogue model representing A. In order to interpret and evaluate the conversational behaviors further, they need to be related to features in A's current internal state. To be able to do this, the incoming conversational behaviors are grouped into Categories by the dialogue model.

The categories are then used to call upon the appropriate *formation* and *updating rules* to form or alter the values of those internal state features that are associated with the perceived conversational behavior. Determining which internal state features can be associated with the conversational behaviors in a particular Category is done based on the *values* of the conversational behaviors' `properties`.

Because the categories specify which internal state features are associated with the conversational behaviors in each category, it is easier for the modeler to construct the *formation* and *updating rules*. This also allows for new conversational behaviors to be easily integrated into the dialogue model, because they don't have to be linked individually to various *formation* and *updating rules* for internal states.

After a conversational behavior is processed, A needs to select and perform an appropriate response behavior in order to further the dialogue. This response behavior is the second perspective from which conversational behaviors can be seen; as the output of the dialogue model. The output of the dialogue model consists of the conversational behaviors that are performed by A. The new configuration of the internal state features that is constructed by processing the input determines from which Category an appropriate response behavior should be selected. The set of categories of output conversational behaviors is the same as the one used to categorize the input conversational behaviors.

To recapitulate, the input of the dialogue model consists of the conversational behaviors plus the descriptions of their `properties` and *`values`*. The properties and value are determined through observation and interpretation and allow for the categorization process of the conversational behaviors by the model. The Categories link the conversational behaviors to *formation* and *updating rules* of the internal state of A. At the output side of the dialogue model, the categories link A's new internal state features to a selected output conversational behavior. These output conversational behaviors receive their properties from the internal state of A.

The conversational behaviors are categorized based on the following selection of their properties:

a) The `intended effect` of the conversational behavior, as aimed for by the speaker.
b) The `effect` of the conversational behavior that is `expected` by the speaker.
c) The `configuration of internal state features` that must hold, in order for the conversational behavior to be selected.
d) The relations between the `behavior type` of the conversational behavior and the behavior types of conversational behaviors possibly performed earlier, by either the speaker or the listener, and to the types of possible response behaviors in the near future.

The other `properties` of the conversational behaviors (i.e. e), f) and g)) are descriptions of the way the conversational behaviors manifest. For example: a question asked in an angry manner, wherein this display of emotion is not explicitly intended, will be grouped in the Gather information category, similar to the same question when it is asked in a neutral manner. These other properties don't contribute anything to the categorization process of the conversational behaviors, but do affect the internal state of A. They are not used to categorize perceived conversational behaviors, but can be called on by the internal state *formation* and *updating rules* directly (See section [6.1.2]).

Categorization of conversational behavior is done in the following manner. After determining the properties and values of the conversational behavior, on both level 1 and level 2 (done for example by annotation as shown in chapter [5]), the values

of the behavior properties are compared to the values of the categories. Each category has a distinctive set of property values that determine which behaviors may be included in that category. This comparison is done for the level 1 values, as more detail about the properties is not necessary for the categorization process. When the values of the conversational behavior match those of a category for at least one of the properties, the conversational behavior can be placed in that category. Thus it is possible for a conversational behavior to be placed in more than one category. As mentioned earlier, we determined a large number of the values of the conversational behaviors performed in the bad news conversation, both on level 1 and level 2. Studying the similarities between the values has led to the construction of the following set of categories:

Behavior categories:

Information seeking behaviors
- Gather information

Information providing behaviors
- Inform about bad situation
- Inform about good situation
- Show agreement
- Show disagreement
- Explain (provide additional information)
- Elaborate

Social management behaviors
- Greeting
- Return greeting
- Giving excuse / apologizing
- Thanking
- Say goodbye
- Return goodbye

Expressing emotions related to oneself behaviors
- Express Joy
- Express Sadness
- Express Fear
- Express Surprise
- Express Confusion
- Express Hope

Expressing emotions related to others behaviors
- Express Empathy
- Express Sympathy
- Express Pity (Sorry-for)
- Express Happy-for
- Express Anger
- Express Praise
- Express Blame

Below we describe of some of the categories, listing the specific values for that category. In addition, we give some examples of the conversational behaviors from the analyzed bad news conversation that are grouped into that category.

Gather information: Typically, conversational behaviors that fall into the Gather information category take the form (i.e. have the behavior type) of *questions*. Alternatively, they can be formulated as directives (e.g. *instruct* or *suggest*) that cause to the listener to provide information.

a) `Intended effect:` A conversational behavior from the Gather information category is intended by A to obtain more information about some aspects of a situation by getting B to provide such information. These kind of conversational behaviors are aimed at possibly getting a (dis-) confirmation of uncertain beliefs, evaluating existing beliefs (by asking for a clarification or explanation) or to form new beliefs about a situation (for example by asking a question such as "can you tell me about your work?"). The communicative function, as described in the DIT++ taxonomy, associated with conversational behaviors in this category is *question*.

b) `Expected effect:` The expected effect of conversational behaviors in the Gather information category is either that B will provide A with the requested information, or that B will initiate a sub-conversation in which he will ask A an question of his own, for clarification.

c) `Configuration of internal state:` In order to select a conversational behavior from the Gather information category, the following feature values must hold in the internal state of the speaker:
   - A has the **intention** to get to know X.
   - A has the **belief** that B knows (i.e. believes) information X.
   - A has the **intention** to obtain information X from B.
   - A has the **belief** that B will provide information X to A.

d) `Relation to the types of other behaviors:` The behavior types most commonly associated with conversational behaviors from the Gather information category are *questions*. The conversational behaviors following those most often have the *statement* behavior type and are aimed at providing information (i.e. the answer) to A. Alternatively; the response behaviors are Gather information behaviors themselves (i.e. counter-questions) and are used to inquire after a clarification or an explanation. There are no typical kinds of conversational behaviors preceding Gather information behaviors, with the exception of an unclear Gather information conversational behavior performed by B.

- Examples of behaviors performed in a bad news conversation that are grouped in the Gather information category are: "What do you mean, spread through the abdomen like mine has?", "What do you mean, invaded?" and "What the hell is going on here?"

Providing information through a conversational behavior can be done in many different ways and can signify a variety of things, such as stating **beliefs** about facts or situations, agreeing or disagreeing with an interlocutor, or explaining a previously performed Inform about good/bad situation behavior by giving additional information. Because of this wide range of possible ways and functions, categorization of conversational behaviors that provide information is spread out over several categories that each represent a different way and function for providing information. All different categories of information providing conversational behaviors are mentioned in the overview list above. As an example, we present the category Inform about bad situation here:

Inform about bad situation: Contained in the Inform about bad situation category are conversational behaviors through which the speaker (A) aims to impart knowledge about a situation that will negatively affect the internal state of the listener (B).

a) `Intended effect:` By performing conversational behaviors from this category, the speaker (A) aims to communicate his beliefs (about the bad situation B is in) to the listener (B) so that afterwards B also believes that he is in a situation that negatively affects him. The conversational behaviors from the Inform about bad situation category have the communicative function *inform* as described by the DIT++ taxonomy.

b) `Expected effect:` The expected effect of an Inform about bad situation conversational behavior depends for a great deal on the beliefs A has about B's attitude towards these kinds of situations and about B's emotion state. In the context of a medical bad news conversation A can expect B to adopt one of the following coping strategies, which are adopted based on B's disposition and his emotions: ***Denial, Anger, Bargaining, Depression*** or ***Acceptance***. The exact effect that A expects his behavior to have on B's internal state depend on which strategy A believes B will select, but it will almost certainly include a stimulation of B's negative emotions.

c) `Configuration of internal state:` The following configuration of internal state features must hold for the conversational behaviors to be included in this category:
   - A has the **belief** B is in a bad situation.
   - A has the **belief** B does not have the belief he is in a bad situation.
   - A has the **intention** to inform B of the bad situation.
   - A has the **belief** he is socially obliged to inform B of the bad situations.

   Note that of these last two internal state features at least one needs to hold before the behavior can be selected.

d) `Relation to other behaviors:` Conversational behaviors that are grouped in the Inform about bad situation category are primarily of the behavior type *statement*. These kinds of conversational behavior can performed without being related to other, previously performed conversational behaviors. But they

can also be used as a response to a conversational behavior which has the communicative function of *question*. In that case the behavior type of the conversational behavior is an *answer*. What the type of the response behavior following an Inform about bad situation behavior is depends on the coping strategy adopted by B.

- Examples of conversational behaviors that fall into the Inform about bad situation category include "The course of chemo didn't cure the tumor.", "The cancer has spread through the abdomen." and "This means the cancer is incurable."

In order to hold a proper dialogue, both interlocutors need to adhere to certain social norms that structure a conversation. Social management behaviors are used to achieve this social structure. The purpose of performing social management behaviors is to increase the social relation that exists between the two interlocutors by affecting the internal state features rapport and liking. It is important to mention that aspects such as politeness or formality that are used in conversational behaviors are not represented in these categories. This is because these aspects only influence the form in which the conversational behaviors is being performed and do not influence the purpose of the conversational behavior. Affecting the internal state features can be achieved through various different kinds of conversational behaviors. Accordingly, several categories are determined in which social management behaviors can be grouped. Given here is the description of the category Giving excuse:

Giving excuse: Conversational behaviors grouped into the Giving excuse category are used to prevent the listener of assigning blame to the speaker for a situation that negatively influences B's internal state. Secondly the conversational behavior is used to convey the speaker's regret about the situation. In the context of bad news conversations behaviors from this category are quite commonplace.

a) Intended effect: The intended effect of a conversational behavior from the Giving excuse category is to make the listener (i.e. B) *form the belief* that the speaker is not accountable for the bad situation or that there are mitigating circumstances to justify his behavior. The indirect intended effect of a Giving excuse behavior is to *redirect or reduce the Blame* that B possibly will assign to A. The communicative function of behaviors in this category, according to the DIT++ taxonomy, is a form of *correction*.

b) Expected effect: The expected effect of a Giving excuse is that B will *assign Blame* to someone or something else than A or, if he assigns blame to A, that this blame is lessened.

c) Configuration of internal state: The following internal state feature values must hold in order for A to be able to select this type of conversational behavior:
   - A has the **belief** B is in a bad situation.
   - A has the **belief** that B believes that A is the cause of the bad situation he (B) is in.

- A has the **belief** that B will assign **Blame** to A for the bad situation.
- A has the **intention** not to receive **Blame**.
- A has the **belief** that information X will divert B's **Blame** to someone else.

d) Relation to other behaviors: Behaviors from the Giving excuse category are often preceded by a behavior from the Inform about bad situation category, which is also performed by A. In those cases, the Giving excuse behavior is often combined with a conversational behavior from theExplanation category. Giving excuse behaviors can also be performed before or after B performs a conversational behavior from the Express blame category. Most often behaviors from the Giving excuse category have the behavior type of *statements*.

- The following sentences are examples of conversational behaviors of the Giving excuse category: "I'm sorry, but I thought you knew." (In this case the last part of the conversational behavior is of the Explanation category.) Another example is "The surgeon could not remove the tumor." (Which is an attempt to direct the Blame to the surgeon).

A considerable part of the conversational behaviors in a bad news conversation is filled by expressions and displays of the interlocutors' emotions. Sometimes displays of emotions are not the purpose of the conversational behavior, but rather a side-effect caused by the strong feelings in the speaker's emotion state. Other times, the expression of the speaker's emotions is meant to communicate those feelings to the interlocutor. Conversational behaviors which contain displays of emotions, but that do not have the purpose to convey an emotion are not included in the following categories, but are categorized based on what their primary intended effect is.

We make a distinction between two sets of categories for conversational behaviors that communicate the speaker's emotions. The distinction is based on whom the emotion expressed in the behavior is directed at. This can either be the speaker himself or someone else, for example the listener or a third person not participating in the conversation. In other words, one set consists of categories that contain conversational behaviors that express emotions related to oneself (i.e. the speaker), while the other set consists of categories that contain conversational behaviors that express emotions related to others. We argue that the following emotions are related to the speaker himself: **Joy, Sadness, Fear, Hope, Surprise** and **Confusion**. The emotions that are more related to a person other than the speaker are **Empathy, Sympathy, Pity (Sorry-for), Happy-for, Anger, Praise** and **Blame**.

As an example of the categories that contain behaviors that express emotions, we present the category Express Empathy:

Express Empathy: By expressing Empathy, A indicates that he understands the negative emotions B is experiencing. This does not entail that A feels the need to do something about the negative emotions, only that he comprehends why B is experiencing them.

a) Intended effect: By performing a conversational behavior that can be grouped in the Express empathy category, A intends B to form the belief that A under-

stands (i.e. has the belief) that the negative emotions B is experiencing are a result of the situation B is in. A also intends to have B form the belief that A understands why B is experiencing those negative emotions. The DIT++ taxonomy does not describe the communicative functions of emotions and subsequently no communicative function is included for an Express Empathy conversational behavior.

b) `Expected effect:` Express Empathy behaviors do not focus on achieving any effect other than informing B of A's understanding of the bad situation B is in. As such, A does not really expect a response of B to his Express Empathy conversational behavior. However, see d) for more information.

c) `Configuration of internal state:` The following configuration of internal state features must hold in order for A to select a conversational behavior from the Express Empathy category.
   - A has the **belief** that B is in a bad situation
   - A has the **belief** that B is experiencing negative emotions.
   - A has the **belief** about how the bad situation and the negative emotions are related. (This belief must have been formed in A's internal state as a result of A being in the same bad situation and experiencing the same negative emotions.)
   - A has the **intention** to inform B about his belief about how the bad situation and the negative emotions are related.
   - A has a high value for the intensity of the emotion **Empathy**.

d) `Relation to other behaviors:` An Express Empathy conversational behavior is often combined with a conversational behavior from the Express Sympathy category. They are either merged into one single conversational behavior or performed consecutively. These two terms (i.e. empathy and sympathy) are often, incorrectly, used as synonyms for each other. Coarsely said: Empathy is the internal experiencing of what another is feeling, but not being motivated to take an alleviating action as a result (i.e. "I know your pain"). Empathy is the 'knowing of another's pain' through one's own experiences. Sympathy is realizing (perhaps cognitively) that someone is upset and desiring to alleviate that, but not experiencing their sense of upset directly and internally as an emotional state yourself. (i.e.: "I imagine your pain (but have not experienced it myself) and wish I could help."). The purpose of using an Express Empathy conversational behavior in combination with a behavior from the Express Sympathy category is to increase the degree of the offered comfort by indicating the hurting is understood.

   Conversational behaviors that contain expressions of Empathy usually take the form of *`statements`*. They are not related to another particular type of behavior. However, Express Empathy conversational behaviors are often preceded by conversational behaviors that contain lots of displays of B's negative emotions.

- Examples of conversational behaviors from the Express Empathy category that occur in a bad news conversation are "I know it's hard to hear this news" and "I understand that this makes you sad."

6.1.2 Behavior properties not used in the categorization

While only four of the properties of a conversational behavior are used to determine in which category the behavior can be grouped, the remaining three properties also play a role in processing the perceived conversational behaviors and selecting and performing the output behaviors. While the first four behavior properties are called upon by the *processing rules* only after the conversational behavior has been categorized, the remaining three behavior properties (i.e. properties e), f) and g)) can be called upon by the *processing rules* directly after the conversational behavior has been interpreted.

For example, contained in property e) are the `emotions` A ascribes to B, based on A's interpretation of the displays in B's conversational behavior. If the expression of emotions is not the primary purpose of the conversational behavior, the ascribed emotions can be used directly by the *belief formation rules* and *belief updating rules* to form new **beliefs** or update existing ones, without having to know into which category the conversational behavior will be grouped.

The same holds for the `social relation` that holds between the two interlocutors. The speaker's perspective of the social relation that holds between the interlocutors is described by property f) and is determined by A's interpretation of the social displays present in the conversational behavior. If the displays of the ascribed social relation do not fulfill the primary function of the conversational behavior, the social relation can be used directly by the *belief formation rules* and *belief updating rules* to form new **beliefs** or update existing ones.

It is interesting to note that the social displays themselves can be used by the *social relation updating rules* to possibly influence the features of A's perspective of the social relation. Examples of such social displays are the degree of politeness that is used in the conversational behavior, the degree of formality of the conversation and the expression of rapport between the two interlocutors.

The third property not used for categorization, i.e. property g), describes the `semantic content` of the conversational behavior. The semantic content is used to describe the meaning of the conversational behavior on the signal level. (See chapter [2] for more information about signal meaning and speaker meaning). The semantic content is used in conjunction with the configuration of the internal state ascribed to B by A, i.e. property c), to determine how to fill the specific second level values of A's internal state features.

In addition to affecting the internal state features of the person addressed by the conversational behaviors, instances of these three properties can also be present in the output behaviors of that person. *`Displays of emotions`* and / or *`expressions of the social relation`* are applied to the conversational behavior after it has been selected.

The selection of which displays of emotion and social relation are to be performed in the conversational behavior is determined by the emotion state-base and the social relation-base. The `semantic content` of the output behaviors (i.e. what actually is said) and the `communicative functions` (i.e. `intended effects`) are determined by other features of the internal state (e.g. the speaker's **beliefs, desires** and **intentions**).

6.2 Constructing the dialogue model

As mentioned in the previous section, the categories of conversational behaviors are used at both the input side and the output side of the dialogue model. When B addresses A in a conversation, the dialogue model representing A receives B's conversational behavior and its described properties as input and subsequently processes that behavior by altering A's internal state. Next, A needs to select and perform an appropriate response behavior. When selecting a response behavior, the set of categories in the dialogue model provides the behavior selection process with conversational behaviors (including their properties) that are appropriate to use as output. In other words, the set of categories, which is contained in a component in the dialogue model called the behaviors-base, functions as storage for all incoming and outgoing conversational behaviors of the dialogue model. In addition, each category calls upon a number of different components of the dialogue model that contain *processing rules* (for example *belief formation rules*, or *intention update rules*) in order to link the properties of the conversational behaviors to A's internal state.

Whenever we talk about conversational behaviors with respect to the input of the dialogue model, we mean B's conversational behaviors that are perceived by A and subsequently interpreted. A's subjective interpretation of the conversational behaviors is contained in the values of the behavior properties. As mentioned in section [6.1], interpreting (i.e. determining the values of the properties of) each conversational behavior in the example dialogue was done via the two analyses presented in chapter [5].

The cognitive processing of the conversational behaviors and the internal state features that are involved in this processing are represented in components that make up the dialogue model. In order to represent the processing in a manner that is cognitively plausible, there are three major aspects that need to be modeled:

- the `properties of the conversational behaviors` performed in the dialogue.
- the internal states of both interlocutors, as viewed by A.
- the mechanisms and *rules* that describe the relations between the conversational behaviors and the internal state features, and between the internal state features themselves.

In this section we give an overview of the components of the dialogue model that evaluate B's conversational behaviors that are perceived by A, contain and modify A's internal state and lead to selection of A's response behaviors. To refresh, A is considered to be the person whose mental model is represented by the dialogue model. Person B is A's interlocutor. In the context of the analyzed bad news conversation, we consider A to be the patient and B to be the doctor.

6.2.1 Overview of the components in the model

The dialogue model consists of various components that interact with each other. These interactions represent the cognitive processing of conversational behaviors and the selection of response behaviors. An overview of the components is presented in figure [6.2]. Detailed descriptions of each of the components are provided below.

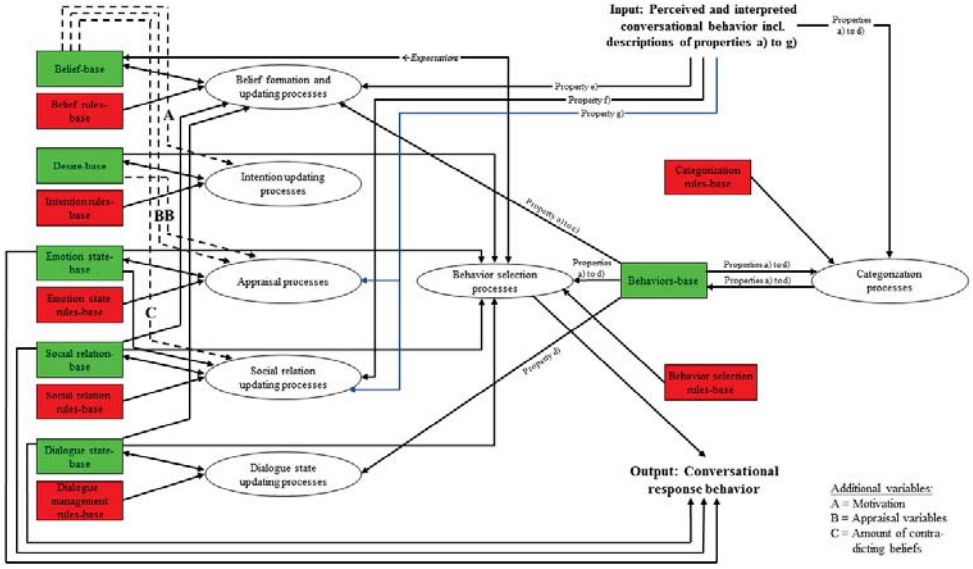

Figure 6.2: Overview of the dialogue model.

The components in the dialogue model are collections of information describing things such as internal state features, processing rules and various processing mechanisms. This information consists of 1) representations of internal state **features** and their associated values and 2) the *rules* that direct the formation of new and updating of current internal state feature values. Each of the components is filled with information before the start of the conversation to allow the dialogue model to function. This includes all the *rules*, A's starting **beliefs, desires** and **intentions**, the starting values of the features that make up A's **emotion** and **social states**, and the state of the dialogue.

We call the representations of the internal state features and their associated values, the elements of the dialogue model. In figure [6.2] a distinction is made between three types of components. Green colored components contain sets of **elements**, red components consist of sets of *formation* and *update rules* and elliptical components represent the mental processes that occur during the processing of the conversational behaviors. These latter components do not have content of their own, but use elements and rules from appropriate other components. In other words, the mental process components in the dialogue model represent the application of rules to elements.

Although the different processes involved in processing of the conversational behavior normally occur simultaneously (e.g. the belief-formation, emotion appraisal, social relation updating, etc), the dialogue model still follows a processing cycle. This means that for example, the formation of new **beliefs** based on the interpretation of the conversational behavior (i.e. the `behavior properties`) and the formation of even newer **beliefs** based on those new **beliefs** happen in two different iterations of the belief formation process. This division into steps is done to allow other processes, such as for example the response behavior selection process, to take the content of a

component at a certain step (or time-frame) and use that particular content.

As mentioned in chapter [4], the basis of the cognitive dialogue model consists of several core components that facilitate cognitive processing. We have taken the folk-psychological model of human practical reasoning by Bratman (1987), also known as the model of **Beliefs, Desires** and **Intentions (BDI)**, as guideline for modeling these cognitive core components. The reason for doing so is that the BDI model is a good representation of how humans think that humans cognitively process events in the world. This allows for a practical and structured representation of the internal state features that are ascribed to the interlocutor (B) within the dialogue model, according to an intuitive and generally shared idea of what these internal state features should be. Furthermore, in computer systems the BDI model is also widely used in the development of intelligent agents, as it allows for the separation between the activities of deliberation on actions (i.e. planning) and the performance of the selected plans (i.e. execution). Using the BDI model as a basis has led to the following core components of the dialogue model.

Behaviors-base: The behaviors-base functions as a repository for all conversational behaviors that are performed or can be performed during the dialogue. All the conversational behaviors in the behaviors-base are grouped into the different categories described in section [6.1.1]. Before the start of the dialogue, the behaviors-base is filled with potential conversational behaviors A might perform during the conversation. During the dialogue new conversational behaviors are added to the behaviors-base via the categorization processes. These new behaviors are the ones performed by B that are perceived and interpreted by A. The interpretation process (which is not included in the dialogue model) determines the properties of the conversational behavior.

By using the *matching function* from the categorization rules-base, the dialogue model determines into which category a conversational behavior can be grouped. This happens in the categorization processes, where the category properties and some of the `behavior properties` of the interpreted conversational behaviors are compared with each other. Only the level 1 values of behavior properties a) to d) are used by the categorization process to be compared with the category properties. The remaining behavior properties (i.e. e), f) and g)) and their values are used by other processes.

In addition to providing input to the categorization processes (in the form of category properties and their *values*), the behaviors-base also provides input to the belief formation and belief updating processes. This input consists of the `behavior properties` and the associated *`values`* of the conversational behaviors that are to be processed. Both level 1 and level 2 properties are used by the belief formation and belief updating processes and thus are provided by the behaviors-base. See the paragraphs explaining the belief-base and the belief rules-base for more information on these processes.

Furthermore, the content of the behaviors-base is also used by the dialogue management processes to monitor and control the progress of the dialogue. The dialogue management processes use the information that is contained in property d) of the conversational behavior that are to be processed, i.e. its behavior type and how its

behavior type is related to those of other conversational behaviors.

In addition to providing input to the processes that alter the current internal state of A, information contained in the behaviors-base is also used during the selection of output behaviors, which are aimed at altering the internal state of B. The behaviors-base is linked with the behavior selection rules-base though the behavior selection processes. The behaviors-base provides the behavior selection processes with possible behaviors that are subsequently evaluated and handled by the *behavior selection rules*.

Categorization rules-base: The categorization rules-base contains *categorization rules* which operate via a matching function. These *categorization rules* are used to compare the *behavior property values* of the observed and interpret conversational behaviors with the *property values* ascribed to each of the categories. As shown in section [6.1], each category contains a set of properties and values that hold for all conversational behaviors gathered in that category. If at least one of the property values of the input behavior matches the values of one of the category properties, that specific conversational behavior can be grouped into that category.

Belief-base: The belief-base contains the collection of **beliefs** A has at any given point. The beliefs in the model represent all the knowledge and assumptions about the situation that A has formed over the course of the conversation plus the beliefs he already possessed at the start. In the BDI model, the term *'belief'* is used instead of knowledge to represent that what an agent believes may not necessarily be true in the environment and that the belief might change over time. We have restricted the beliefs in the belief-base to only be related to the context of the conversation in order to keep the model clear and understandable.

The belief-base contains three types of **beliefs** that A can have. Most prominent are the beliefs A has about the state of his current situation, which includes features in his own internal state. Secondly, the belief-base contains special beliefs about the anticipated outcomes of the conversational behaviors A performs, called **expectations**. Thirdly, the belief-base also contains A's beliefs (i.e. assumptions) about B's internal state. This includes assumptions about B's beliefs (which are also divided into the three types), B's desires/goals, B's intentions and B's emotions. Each belief in the belief-base receives a certainty value that represents how strongly A believes that certain belief to be true in the situation. This certainty value is determined based on social values of trust, liking and rapport A has with respect to the source of the belief. More information on how belief certainty values are determined can be found in section [6.2.3].

Content of the belief-base is generated by the belief formation and updating processes. The belief formation processes form new beliefs by applying *belief formation rules* to the information provided by both the behaviors-base and the interpreted conversational behavior itself. Alternatively, the *belief formation rules* can construct new beliefs from beliefs already present in the belief-base. In addition to the belief formation processes, belief updating processes can alter the content of the belief-based by applying *belief updating rules* to current beliefs in the belief-base. These current beliefs are then updated into new beliefs.

In addition to the beliefs, the belief-base also contains a special list that consists of pairs of contradicting beliefs. These contradicting belief-pairs are used to ensure that the beliefs in A's belief-base are consistent with each other and that no contradicting beliefs can occur at the same time in the belief-base. Whether or not two beliefs contradict depends on the content of the conversational behaviors from which the beliefs are formed. The cognitive mechanisms that determine which beliefs contradict each other are not included in the dialogue model, because these mechanisms rely on the process that interprets the semantic contents of the conversational behaviors. The interpretation of the properties of conversational behavior, especially the semantic content, is a complex cognitive process that is still rather unclear.

As a result no new contradicting beliefs-pairs can be constructed for newly-formed beliefs in the dialogue model during the conversation. It is important to mention that the list of contradicting belief-pairs is not a proper representation of the knowledge in the internal state of A, but rather a practical way of representing A's beliefs about the contradictions. An alternative way of representing knowledge about contradictions would be to include a single belief in the belief-base for every contradiction A knows.

Belief rules-base: The belief rules-base component contains two types of rules. The first type of rules, called *belief formation rules*, enables the construction of new beliefs. New beliefs can be formed based on the properties of an interpreted conversational behavior or alternatively based on beliefs currently present in the belief-base.

When beliefs are formed based on the properties of a conversational behavior, the input of the *belief formation rules* is provided by the behavior base in the form of properties a) to c) (i.e. the intended effect and the configuration of the internal state underlying B's conversational behavior) and directly by the conversational behavior itself through property e) (i.e. B's emotions). Often these new beliefs are assumptions about features in B's internal state that are associated with that specific conversational behavior. For example, A forms a belief about what intention B had prior to or during the performance of his conversational behavior, based on the intended effect property A has interpreted in B's behavior. Another example is that A can form a belief about what B is feeling, based on his (A's) observation and interpretation of a display of emotion in B's conversational behavior (described by property e)).

Beliefs that are constructed based on the beliefs currently present in the belief-base can be the result of A adopting B's internal state features as his own. For example, if A strongly believes (i.e. has a belief with a high certainty value) that a belief of B holds, A might adopt that belief as one of his own beliefs. Imagine that B says "The tumor has spread through your body". One of the internal state feature ascribed to B by A when he categorizes that conversational behavior is that B has the belief 'The tumor has spread through A's body'. A will form a belief about this: A believes that B believes 'The tumor has spread through A's body'. If A believes this fact strongly enough, A might adopt this belief also: A believes 'The tumor has spread through my body.' Alternatively, new beliefs that are formed on current beliefs can be the result of characteristics such as causality or transitive relations.

Determining the certainty values of new beliefs that are based on the conversational behaviors of B is strongly influenced by the elements in the social relation-base.

If the value of the **trust** A has with respect to B is high (i.e. A trusts B a lot), then A will often assign higher certainty values to the beliefs that are formed based on behaviors performed by B.

The second type are called *belief updating rules*. Unsurprisingly, these types of rules are used to alter beliefs currently present in the belief-base. In the dialogue model, we make a distinction between two kinds of updating mechanisms and each has its own set of rules.

On the one hand there are *belief updating rules* that ascertain that newly formed beliefs do not contradict beliefs that are already in the belief-base. These rules operate by checking whether the new belief matches one of the beliefs in the lists of contradicting belief-pairs. If this is the case, then the *belief updating rules* will delete the belief of the belief-pair that has the lowest certainty value. However, if both beliefs have low certainty values, then both beliefs are kept in the belief-base.

On the other hand, there are *belief updating rules* that make small changes to current beliefs, such as an increase or decrease in certainty. Alternatively, the update of a belief can be a small adjustment to the value of a belief, such as A adjusting his belief about the intensity of one of B's emotions.

In addition to the *belief formation rules* and the *belief updating rules*, the Belief rules-base also contains rules (called *strategy rules*) that represent the particular coping strategies or protocols, which are discussed in chapters [3] and [4]. These *strategy rules* operate as extra filters on the outcomes of the other belief rules.

Let us assume that the dialogue model follows the ***denial*** strategy of coping. When the dialogue model forms a new belief based on what the interlocutor is, the strategy rules flip the truth-value on the new belief immediately. For example, when B says "you're sick" and A might normally form the new belief 'I'm sick', A instead forms the new belief 'I'm not sick.'

Desire-base: Contained in the desire-base are all the desires A has at a certain point. This set of desires represents all the possible states of a situation or a person that A wants to be the case. In a conversation, most desires focus on bringing about a particular configuration of the internal state of the interlocutor and subsequently causing them to perform a conversational behavior that is associated with that configuration. As a result, they perform conversational behaviors that may result in the desire being fulfilled. However, people do not attempt to fulfill all desires they have at the same time, but single out the desire that is most appropriate at that time and commit to it. The commitment to a desire, i.e. the dedication to perform conversational behaviors to see the desire fulfilled, causes that **desire** to turn into an **intention**. As a result, at any one time during the conversation the desire-base in the dialogue model will contain a number of desires and a single intention.

As it lay outside the scope of our model, we studied the process of desire formation / adoption only in a general sense and did not include the possibility to create new desires during the conversation in the dialogue model. In our iteration of the dialogue model as a working example, we have included a set of desires that are derived from the analyses performed in chapter [5]. Based on the analyses, we were able to determine different types of desires that may hold during bad news conversations.

The following types of desires are found:

- Desires to get information from the interlocutor. (about the situation or about the internal state of the interlocutor)
- Desires to make an alteration in the internal state of the interlocutor. (e.g. to elicit an emotion)
- Desires to manage the dialogue.
- Desires to manage the social relation between the two interlocutors.
- Desires to express something from one's own internal state.

Intention rules-base: In order to commit to a desire or to switch commitment from one desire to another, the model employs *intention updating rules*. These *intention updating rules* are contained in the intention rules-base, which can be accessed by the intention updating processes. The intention updating processes function by applying the *intention updating rules* to the set of desires in the desire-base. In addition, the intention updating processes use beliefs from the Belief-base to determine which desire is the most appropriate one to commit to at a given moment. If more than one desire seems appropriate, the intention updating processes can cause the dialogue model to commit to the desire that is the easiest to obtain.

Just like in the Belief rules-base, the Intention rules-based also contains *strategy rules*. The strategy rules prevent the dialogue model to commit to desires that are not appropriate given the coping strategy or protocol the dialogue model is following.

Emotion state-base: The emotion state-base represents how A is feeling emotionally with respect to his own situation and with respect to the situation his interlocutor is in (including the interlocutor's internal state). Included in the emotion state-base are all the different kinds of emotion A can possibly experience during the conversation. These different kinds of emotion are grouped in a fixed set, which consists of the 22 emotions types presented in the OCC model (Ortony et al., 1988) plus additional emotion types for the emotions Surprise, Confusion, Sympathy and Empathy. Each element in the emotion state-base consists of an emotion type (which is fixed) and a corresponding intensity value (which is variable). Alteration of these elements during the conversation is the result of performing the emotion appraisal processes. The emotion appraisal processes increase or decrease the intensity values of the emotions depending on which rules from the emotion state rules-base are being applied by the processes.

In addition to the appraisal processes, some of the elements in the emotion state-base can be influenced by the coping strategy the dialogue model is using. In chapter [3] we describe Kübler-Ross' five stage theory. Two of the five stages (i.e. Anger and Depression) indicate that ***displaying emotions*** is a way of coping with bad news. The way these coping strategies are represented is by a onetime, fixed increase in the intensity value of either the emotion type **Anger** or the emotion type **Sadness**.

In addition to storing the emotion types and the intensity values that A is experiencing, the emotion state-base also provides information that can affect the output conversational behaviors. The output behaviors can be affected by the elements from the emotion state-base in two different ways. On the one hand, emotions can play an

important role in the selection process of the output behavior. This occurs when one or more of the preconditions of a conversational behavior needs to be fulfilled by an emotion. Moreover, when the intensity value of an emotion is very high, the selection process can be overridden. In that case a conversational behavior from the Express emotion category is automatically selected, expressing the emotion that possesses the high intensity value.

Besides influencing the process of behavior selection, the emotions can also affect the manner in which the selected output behavior is being performed. This is done by altering the form of the output behavior (e.g. by altering the prosody) or including displays and expressions of the relevant emotions. The process of constructing the precise form of a conversational behavior (i.e. the behavior realization process) was not included in the dialogue model, because the focus of the dialogue model lies on the behavior processing and selection processes and not on the realization process. When looking at embodied dialogue systems, often the behavior processing / selecting modules are separated from the behavior realization modules due to their different functions.

Emotion state rules-base: The rules that are used by the emotion appraisal processes to update the elements in the emotion state-base are called *emotion appraisal rules* or *emotion state updating rules* and are contained in the emotion state rules-base. These rules are based on some of the theories about emotion appraisal discussed in chapter [4]. More specifically, the *emotion appraisal rules* in the dialogue model are constructed by using a selection of the variables used in the OCC model (Ortony et al., 1988), Clark Elliot's Affective Reasoner model (Elliot, 1992) and the Émile system (Gratch, 2000). For a complete overview of the appraisal variables used in the dialogue model, see the detailed explanation of the emotion appraisal rules in section [6.2.3].

The *emotion appraisal rules* evaluate how A's interpretation of the situation expressed in B's conversational behavior relates to the subjective way that A perceives the world. Alternatively, the *emotion appraisal rules* can also be used by the appraisal processes to evaluate how A's view of the world relates to changes in the internal state that are not directly caused by a conversational behavior performed by B, for example, the formation of a new belief or A altering his intention. Depending on the results of the evaluation of the internal state elements and behavior properties of B's conversational behavior, the *emotion appraisal rules* will either increase or decrease the intensity values of the associated emotions.

The way that A perceives and feels about the world is called A's **disposition**. A's disposition consists of his desires and intentions to bring about certain situations (e.g. the desire to cause a change in the internal state of B or the desire for B to perform a particular type of conversational behavior), the social relation A has with respect to B (i.e. his social disposition) and traits of A's personality (e.g. whether A is hotheaded or stubborn). Personality traits are not included in the dialogue model as they are outside the scope of the study.

In the model, A's disposition influences the manner in which he interprets the situations expressed in B's conversational behaviors. For example, whether the inter-

preted situation is positive or negative for A depends on his desires and/or whether or not he believes what B is saying depends on A's trust in B. Gratch et al. (2006a) call the interpretation of the relation between the situation and the disposition the "causal interpretation".

As a result of this complex network of interdependencies between various types of internal state elements, the emotion appraisal processes are linked to the belief-base, the desire-base, the emotion state-base, the emotion state rules-base, the social relation-base and property g) of the interpreted conversational behavior (i.e. the semantic content).

Social relation-base: Contained within the social relation-base are internal state elements that, when taken together, represent the social relation (or social disposition) A has with respect to B during the conversation. Similar to the elements in the emotion state-base, the elements in the social relation-base consist of a social aspect type (which is fixed) and an associated intensity value (which is variable). The representation of the social relation in the dialogue model is based on the following social aspect types: the amount of **trust**, the amount of **liking** and the amount of **rapport**, all of which can be influenced by the social roles the interlocutors have adopted. The selection of these internal state features is inspired by various studies, such as (Falcone and Castelfranchi, 2001) for the social aspect type **trust**, and (Gratch et al., 2006b) for the notion of **rapport**. Although the social aspect type **liking** is often used to indicate how a user perceives an agent system, we argue that it can also be used to indicate the perception of the user by the agent. However, the inclusion of **liking** is based more on an intuitive notion than on examined studies.

An example of how the social features can be influenced by the role of the interlocutors is as follows: a patient will in general be more inclined to trust in what a doctor is saying where his medical state is concerned (due to the doctor's expertise) than in the semantic content of the conversational behaviors performed by a close friend. This holds even if his liking of the close friend is higher than his liking of the doctor.

Similar to the emotion state-base, the social relation-base does not only store the social elements of A's internal state, but the information contained in the component can also be used to affect conversational output behaviors. This can be achieved in two different ways. In one way, elements from the social relation-base can be needed to fulfill one or more premises that determine the configuration of the internal state which is needed to select a conversational response behavior. In such a case, the social elements assist in selecting the appropriate response behavior. The other way social elements can affect conversational output behaviors is by altering the form of the output behavior. Most of the time this expresses itself by changes to the prosody, particularly the speech rate and turn switches, but it can also be expressed through gestures such as touching the listener and closer proximity. Note that the process of altering the form of the output behavior is not included in the cognitive dialogue model, but can be placed in a model of the behavior realization.

Social relation rules-base: The social relation rules-base contains the rules that facilitate updating of the social elements in the social relation-base. Applying the *social relation updating rules* to the social elements is done through the social relation updating processes. In addition to the elements of the social relation-base, the *social relation updating rules* require various other elements present in the dialogue model to operate. More specifically, the social relation updating processes retrieve information from the belief-base, the emotion state-base and property g) from the interpreted conversational behaviors (i.e. the semantic content). In addition, the social displays in the conversational behavior (that, when interpreted, lead to property f), the social relation according to the speaker) are also requested by the social relation updating processes.

The *social relation updating rules* evaluate this information with respect to certain standards. For example, if the emotions towards B are negative or B's previous conversational behavior contains displays of impoliteness, the social relation will be influenced negatively. These standards are determined by the social roles of both interlocutors during the conversation. Based on the outcome of the evaluation, the intensity values associated with the various social aspect types of the social relation are either increased or decreased.

Dialogue state-base: The dialogue state-base contains A's interpretation of the state of the conversation at any moment. It does so by keeping track of the types of conversational behavior that have previously been used in the conversation. The `behavior type` of each conversational behavior is stored in a sequenced list in order to keep track of the sequence the behaviors are performed in. The behavior types also indicate which phase the conversation is in, i.e. the opening (characterized by the behavior types *`greeting`* and *`return greeting`*, which are both *`expressives`*), the body (which can contain conversational behaviors of all types) and the closing (characterized by the *`expressives: say goodbye`* and *`return goodbye`*).

In addition to storing the behavior type of each performed conversational behavior, the dialogue state-base is used by the dialogue model to manage the course of the conversation. The management of a conversation consists of two parts. On the one hand, the dialogue management focuses on the selection of appropriate response behaviors given the behavior type of the preceding conversational behavior. For example, if the conversational behavior performed by B has `behavior type` *`question`*, it is most common if the response behavior has `behavior type` *`answer`*. It is likely that a response behavior of that type is most appropriate in the situation.

The second part of the management of dialogues focuses on processes that monitor and control the flow of the conversation. This consists of, amongst other things, the regulation of turn taking, scheduling the conversational behaviors (i.e. planning the timing of performances), the handling of interruptions and so on. This aspect of dialogue modeling is not extensively studied as it falls outside the scope of this thesis. Consequently, it is not included in the dialogue model.

Dialogue management rules-base: The dialogue management rules-base contains rules that update the content of the dialogue state-base and control the management of

the conversation. These *dialogue management rules* are used by the dialogue state updating process to update the sequenced list with the behavior type of the newly interpreted conversational behavior. Information about which behavior types the performed conversational behavior has is provided to the dialogue state updating processes by the behavior-base in the form of property d). Property d) also contains information on how the behavior type of the conversational behavior is related to other behavior types (for example by adjacency pairs). Linking two conversational behaviors based on their behavior types is also done via the *dialogue management rules*.

Note that currently there are no *dialogue management rules* included in the dialogue model that control the governing of temporal aspects of performing conversational behaviors such as turn taking mechanisms, interruption handling, topic shifts etc. However, such rules can easily be included in this rules-base.

Behavior selection rules-base: The *behavior selection rules* contained in the behavior selection rules-base are comparable to those in the categorization rule-base. The main difference is that the *behavior selection rules* operate in reverse compared to the *categorization rules*. The behavior selection process operates by applying the *behavior selection rules* to the current internal state elements of A. The *behavior selection rules* check whether the elements match the premises that hold for each conversational behavior in the behaviors-base.

In addition, the *behavior selection rules* use information from the dialogue state-base about what kind of `behavior type` is most appropriate for the response behavior to have. This narrows down the categories from which a response behavior can be selected as some behavior types are more associated with behaviors from particular categories, as described by property d). For example, conversational behaviors with the *`question`* `behavior type` often fall within the Gather information category.

Also included in the behavior selection rules-base is the rule that causes the selection of an appropriate response behavior (on the basis of the other behavior selection rules) to be replaced by a behavior from the Express emotion categories. This rule fires if the intensity value of the associated emotion is very high. This information is provided by the emotion state-base

It is important to note that the *behavior selection rules* do not determine the manifestation form of the conversational behavior (e.g. displays of emotion, social features such as politeness level, or prosody), but only determine from which category the response behavior is selected.

6.2.2 Representation of the elements

In this subsection, we present the way in which the elements of the components are specified in the dialogue model. Because the elements are representations of the internal state features of A, it is important to come up with a structured and transparent way to express them. The BDI model is frequently used in the development of intelligent software agents and research in that area has led to formal representations of the BDI features. This is done through axiomatization of certain BDI relations (e.g. Belief consistency or Intention-goal compatibility) and expression of BDI in multiple-modal

logic, wherein Beliefs, Desires and Intentions are represented as different modalities (e.g. see Cohen and Levesque (1990)). The multiple-modal representation of BDI was also combined with temporal logic (e.g. see Rao and Georgeff (1991) or Wooldridge (2000)).

Because the BDI features are explored so thoroughly in other studies, the elements that represent the BDI features of the internal state in our dialogue model are expressed in a manner as is used in the formal logic models. Note that our research does not focus on the formal logic representations of BDI features. Consequently, logic operators and functions are not included in the dialogue model.

Elements that represent internal state features outside the BDI core, such as emotions and features that make up the social relation between the two interlocutors are expressed among the same lines as the BDI elements. This is done to maintain continuity in the representation of the internal state features and make it easier to handle elements together in the various processes.

Beliefs: As mentioned in the description of the belief-base (in section [6.2.1]), we make a distinction between three different types of beliefs in the dialogue model. The first type of beliefs conveys what A believes about the state of the situation he is in through the beliefs' atomic values. This type of beliefs is expressed in a quite straightforward manner:

- $BEL_A(x)$, α.

The attribute BEL stands for **belief** and the subscript A indicates that the belief is held by A (this seems redundant, but will become very useful when we start talking about beliefs about beliefs about desires of others etc.). The realization of variable x is determined by the semantic content of the belief and expresses A's interpretation of the situation. As we will show later on, x can also express elements of the internal state of B. Finally, α is the variable that expresses the certainty value of the belief, i.e. the conviction that A has of the belief being true. The value of the certainty can range from 0 to 1, where 0 represents a complete uncertainty and 1 an utter certainty. Note that having a certainty value of 0 does not mean A does not know x, but rather that he has no idea whether x is true. Determining what the specific certainty value associated with a belief should be is very difficult. In the model we determine the certainty values based on educated guesses and comparison to certainty values of other beliefs. An example of one of A's beliefs, if expressed in natural language might be: "I (i.e. A) am quite sure about the fact that my disease is cured". Expression of this belief in the dialogue model would be:

- BEL_A(disease is cured), 0.7.

The second type of beliefs deals with A's thoughts about what his interlocutor B might be experiencing. In other words, they are the beliefs that A forms about B's beliefs, desires, intentions and emotions. These beliefs are often formed based on the perception and interpretation of the conversational behaviors performed by B. We distinguish four forms of expression of this type of beliefs:

- $BEL_A(BEL_B(x), \beta), \alpha$. For beliefs about B's beliefs. The variable β is used here to indicate the certainty value of B's belief x according to A. The certainty value of A's belief is expressed by the variable α.
- $BEL_A(DES_B(x)), \alpha$. For beliefs about B's desires.
- $BEL_A(INT_B(x)), \alpha$. For beliefs about B's intentions.
- $BEL_A(EMO_B(e), \epsilon), \alpha$. For beliefs about one of B's emotions. The variable e expresses which of B's emotions the belief is about. The variable ϵ expresses the intensity value that A believes the relevant emotion possesses.

These are the beliefs of A about the contents of B's internal state that A has derived from B's conversational behaviors. Note that these are not necessarily the features that are actually present in the internal state of B.

The third type of beliefs covers A's expectations about what effects his conversational behaviors will have on his interlocutor. Most of the time, the conversational behavior performed by A will cause some form of behavior to be performed by B. **Beliefs (expectations)** about what the response behavior is, are expressed as:

- $BEL(exp)_A(y), \alpha$.

The basic attribute BEL is extended with the (exp) extension to signify that it is an expectation rather than a regular belief. In the case of expectations, the variable y can be replaced with the following: $ToPerform_B(z)$. The ToPerform attribute indicates that A expects B is going to do something (e.g. perform a conversational behavior) as a response to the conversational behavior performed by A. The variable associated with the ToPerform attribute (i.e. z) expresses what it is that A actually expects B to do as a response. The value of this variable can be a conversational behavior, such as Express Sadness or Show agreement or simply a `type of behavior`, e.g. *`answer`* or *`question`*.

Suppose A poses a question to B about the state of his disease and A strongly anticipates that B will provide him with an answer and A is rather optimistic, then the associated expectations are expressed as:

- $BEL(exp)_A(ToPerform_B(BEH(answer)))$, 0.7
- $BEL(exp)_A(ToPerform_B(\text{Inform about good situation (disease is cured)}))$, 0.7.

If A expects that B will make a change in his internal state as an effect of the conversational behavior performed by A, then the expressions of the second and third types of beliefs can be combined. Instead of the ToPerform attribute and its variable, the variable x of the $BEL(exp)_A$ is replaced with a description of the internal state feature that B is expected (by A) to form. For example, if A tells B he is feeling good even though he is sick, A might weakly expect B to form a belief about this:

- $BEL(exp)_A(BEL_B(\text{A is feeling good}), 0.7), 0.4$.

The certainty value of an expectation (i.e. α) represents the degree in which A believes that the situation can be brought about or the action to occur. In other words a high value for α indicates that (A believes) it is very likely the situation will become true and a low value indicates that the situation is unattainable. Such expectations are formed after a conversational behavior has been selected and performed by A.

Desires and Intentions: Each element in the desire-base represents a desire that is part of A's internal state from the start of the conversation. These desire elements are expressed in the dialogue model as:

- $DES_A(x)$

where the attribute DES stands for **desire** and x is the variable that expresses a state in the environment or the situation that A wants to bring about. In conversations, desires tend to focus on causing the listener to perform a certain type of behavior or focusing on changing some of the internal state features of the listener. Filling in the x variable with a $ToPerform_B(y)$ operator-value pair represents A's desire to get B to perform a certain `type of conversational behavior`, which is specified by the variable y.

Alternatively, when A desires to alter something in B's internal state, the variable x can take the form of an element of B's internal state. For example, if A wants B to know something, A can have $DES_A(BEL_B(y), \beta)$. In this case, the variable y specifies what it is that A wants B to know. This nesting of internal state features can become confusing when several layers are expressed. For example, underlying the sentence "I think you want to be cured" performed by A (i.e. I is A) is the **desire** that A wants B to know (i.e. believes) that A thinks (i.e. believes) B wants (i.e. has the desire) the situation 'to be cured' to be the case. This desire is expressed in the dialogue model as:

- $DES_A(BEL_B(BEL_A(DES_B(\text{to be cured})), \alpha), \beta)$.

When A commits to a particular desire it turns that desire into an intention. The expression of an intention in the dialogue model is the same as for desires, with the exception that the primary attribute is not DES_A for **desire**, but INT_A for **intention**. This results in the inclusion of $INT_A(x)$ elements in the desire-base. Similar to a desire, the value of the x variable of an $INT_A(x)$ element can be filled in by a $ToPerform_B(y)$ operator-value pair or with an element of B's internal state.

Other values that can be used to fill in the variable x are the operator-value pairs $Increase_B(z)$ and $Decrease_B(z)$. These operator-value pairs are used to represent A's intention to change either the emotion state or the social state of B. The variable z can be filled by social or emotion elements so that the conversational behavior that is to be selected will cause the value (i.e. ϵ, τ, λ or ρ) associated with the correct attribute, to be altered.

Emotions: The elements that make up the emotion state-base represent the emotions A may experience during the conversation. In the dialogue model these elements are expressed as follows:

- $EMO_A(e), \epsilon$.

Here, the attribute EMO stands for **emotion**, while the subscript indicates who is experiencing the emotion. The variable e indicates which type of emotion specified in the emotion state base A is experiencing. The emotion state-base contains the

22 emotion types presented in the OCC model (Ortony et al., 1988) plus the emotion types Surprise, Confusion, Sympathy or Empathy. All 26 emotions types are expressed in the emotion state-base. Whether or not A is experiencing them depends on the variable ϵ.

The variable ϵ is used to express the intensity value of an experienced emotion. The intensity value can vary between 0 and 1, where 0 expresses that A is not experiencing the emotion at all and 1 expresses a very strong feeling of the emotion. It is entirely possible for a person to experience different emotions at the same time. If this is the case then multiple emotion types in the emotion state-base have an intensity value higher than 0. Note that this does not mean that the emotions will necessarily be displayed in A's conversational behaviors. Before the start of the conversation, the intensities are arbitrarily set to 0 in the dialogue model. Alternatively, the starting values may vary to indicate a prior emotion state.

Social relation: The social relation that exists between two interlocutors is determined by various internal state features, such as **liking**, **trust** and **rapport**. These internal state features are represented by elements from the social relation-base, each of which consists of a social aspect type (that describes which feature is meant) and an intensity value. The social relation elements in the dialogue model are expressed in the following way:

- TRUST_{AB}, τ. This expresses the **trust** A has in B.
- $\text{LIKING}_{\text{AB}}$, λ. This expresses the **liking** A has of B.
- $\text{RAPPORT}_{\text{AB}}$, ρ. This expresses the **rapport** that exists between A and B.

The TRUST, LIKING and RAPPORT attributes express the social aspect type of the element, while τ, λ and ρ are variables that express the intensity values of the respective social aspect types. The intensity values can all vary between 0 to 1. Similarly to the intensity values associated with the emotions, the values for each of the social aspect types are set to 0 before the start of the conversation. These starting values may vary accordingly to indicate a social relation that is already in place at the start of the conversation.

The subscripted variables operate slightly different than for other elements (e.g. BEL, DES or EMO). Instead of using a single letter to indicate to whom the element belongs, the social relation elements have two variables that indicate the direction of the social relation depending on the order of the variables. For example, the element TRUST_{AB} expresses the trust A has in B.

As the social relation elements are part of the dialogue model that represents A's internal state, the relations are represented as they are perceived by A. Thus, the social relation element TRUST_{BA} expresses the trust B has in A, according to A. An element that expresses a social relation originating from someone else than A may for example be used in $\text{BEL}_{\text{A}}(\text{TRUST}_{\text{BA}}, 0.8), 0.7$ which indicates that A believes that B trusts him quite strongly.

Dialogue state: The elements in the dialogue state-base are used by the dialogue model to keep track of the course of the conversation. It does so by storing the

behavior types of the previously perceived and performed conversational behaviors in a sequenced list. The elements of that list are expressed as follows:

- $\texttt{Performed}_{\texttt{A}}\texttt{(BEH(type))}$

The `Performed` attribute indicates that the `behavior type` belongs to a conversational behavior that has already been performed and the subscript variable expresses who performed the conversational behavior. The attribute BEH stands for `behavior type`, while the variable 'type' expresses the actual behavior type of the conversational behavior. The structure of the dialogue state is not actually represented by the elements themselves, but rather by the order in which they are stored in the sequenced list. The list as a whole represents the flow of the conversation and the interactions between the interlocutors.

In addition to storing the behavior types of the conversational behaviors previously performed in the conversation, the dialogue state-base also contains elements that express the behavior types of the most appropriate conversational behaviors A could perform in response to B's most recent behavior. These suggested behavior types are expressed in the following way:

- $\texttt{ToPerform}_{\texttt{A}}\texttt{(BEH(type))}$

The `ToPerform` attribute indicates that conversational behavior associate with the suggested behavior type, expressed by the attribute-variable pair BEH(type), is yet to be performed by A. Determining which behavior types are most appropriate for the response behavior to have is done by applying the *dialogue management rules* to the most recent `behavior type` in the sequenced list. As mentioned in the descriptions of the belief elements, the ToPerform element can also be used as a value of the variable in one of A's expectations. Note that in that case the subscripted variable changes from A to B, as B will be the one who is going to perform the conversational behavior.

Conversational behaviors: The behaviors-base in the dialogue model functions as a repository for two groups of conversational behaviors. On the one hand, it contains all conversational behaviors that have been performed by B up till a certain moment and on the other hand, it is filled before the start of the conversation with potential conversational behaviors A might perform during the conversation. Furthermore, the conversational behaviors from both groups are categorized based on similarities found in their properties.

All conversational behaviors in the behaviors-base are represented by sentences that describe the utterances performed, plus the representations of properties a) to d) of the corresponding conversational behavior. The properties of the conversational behaviors are expressed in the following way:

- **IntEff(x)** — This expresses the `intended effect` of the conversational behavior (i.e. property a).
- **ExpEff(y)** — This expresses the `expected effect` of the conversational behavior (i.e. property b).

- **Various elements** This expresses the `configuration of internal state features` underlying the conversational behavior (i.e. property c).

- **BEH(type), $\leftarrow$(t) and $\rightarrow$(t)** This express the `behavior type` of the conversational behavior, plus the operators that indicate how the `behavior type` is related to other `behavior types` (i.e. property d).

It is important to note that, with the exception of property c), these representations do not express features of either A or B's internal state, but rather describe the conversational behaviors they perform. As such they do not have a subscript variable. However, properties a), b) and d) are still closely related to particular internal state features as can be seen in the cognitive processes that link the behavior-base to other components in the dialogue model. For example, the IntEff(x) can be linked to an $INT_A(x)$ element, while the ExpEff(y) can be linked to a $BEL(exp)_A(y)$, α element. The `behavior type` of the conversational behavior is used to keep track of the dialogue state.

The variable x of the `intended effect` expresses one of the communicative functions mentioned in chapter [5], plus its associated level 2 value. For example, one of the intended effects ascribed to the conversational behavior "How are you doing?" is to get a response behavior from A that contains an answer to the question. The intended effect (i.e. the communicative function) is represented as follows: IntEff(question(status of A)). Note that the term question here indicates the communicative function of the conversational behavior, not its behavior type. The **intention** that the listener (A) ascribes to the speaker (B) based on this conversational behavior is:

- $INT_B(ToPerform_A(BEH(answer(status\ of\ A))))$.

In other words, B wants (i.e. has the **intention** to get) A to perform a conversational behavior that has the `behavior type` *`answer`*. Such an answering response would then have the intended effect to 'inform B about A's status', i.e. IntEff(inform(status of A)).

The elements described in property c) are expressed in the exact same ways as they are represented in their associated bases, using the same attributes and variables (e.g. **beliefs** are expressed as $BEL_A(x)$, α, **emotions** are expressed as $EMO_A(e)$, ϵ, etc.). The values of the subscript variables of these elements depend on who performed the conversational behavior. For example, if B performed the conversational behavior, then the elements expressing the underlying internal state features are also B's.

The $\leftarrow$(t) and $\rightarrow$(t) operators in property d) can be used to indicate how the `behavior type` of the conversational behavior (i.e. `BEH(type)`) is related to other types of conversational behaviors. These operators are used by the dialogue state updating processes to determine which behavior types are most appropriate for selecting a response behavior (i.e. the $ToPerform_A(BEH(type))$ elements in the dialogue state-base). The $\leftarrow$ operator indicates what the behavior types of preceding conversational behaviors could have been, while the $\rightarrow$ operator indicates what kind of behavior

types are most appropriate for the response behavior to have. The variable (t) can be filled by a representation of a behavior type, i.e. `BEH(type)`.

Non-categorization properties: Properties of the conversational behaviors that are not used by the categorization processes to categorize the behaviors in the behaviors-base may be used by other processes to affect the internal state of the listener (i.e. A). Only the properties of conversational behaviors performed by B are used in the dialogue model. As a result, the elements expressed by these properties are representations of internal state features that A ascribes to B. These internal state features are ascribed to B based on A's interpretation of B's conversational behaviors.

Properties e) and f) of a conversational behavior describe the features that respectively make up B's emotion state and social relation state. These properties are based on interpretation of the displays expressed in the conversational behavior. Consequently, the features are expressed in the dialogue model the same way as corresponding features in the internal state of A. In other words, the emotions of B that are described by property e) are represented as:

- $EMO_B(e)$, ϵ.

and the social elements that make up B's social relation, which are described by property f), are represented as:

- $TRUST_{BA}$, τ. This expresses the **trust** B has in A.
- $LIKING_{BA}$, λ. This expresses the **liking** B has of A.
- $RAPPORT_{BA}$, ρ. This expresses the **rapport** that exists between B and A.

Expressed in property g) is the `semantic content` of the conversational behavior. The representation of the semantic content basically describes the situation / state of the environment that is expressed by the conversational behavior. Contained in this description are objects, propositions, events and so forth that are determined by A's interpretation of B's conversational behavior. The description of the situation is used to fill in the level 2 values of elements in A's internal state. For example, it can be used as a value for the x variables in beliefs and desires in A's internal state.

Because each conversational behavior has specific semantic content, there is no general representation form that expresses the semantic content. Instead the semantic content is expressed as small fragments of the sentence used in the conversational behavior. For example, a conversational behavior in which the sentences "John has evidence of metastases. I know the surgeon has talked with John, but I don't think he fully understands his condition." are spoken, contains the following facts of the semantic content:

- John has evidence of metastases, (i.e. John's situation is bad)
- The surgeon has talked with John.
- John does not understand his condition.

6.2.3 Functioning of the rules

After interpreting a conversational behavior performed by the interlocutor and determining its properties, a person deals with the conversational behavior via various cognitive processes. These cognitive processes are represented in the dialogue model as actions that entail taking elements from the appropriate components and subsequently applying rules from the various rules-bases to these elements. As a result, new elements are constructed and placed in the appropriate components or elements already contained in the components are altered in some way. For example, the belief formation and updating processes take certain elements from the behaviors-base (i.e. properties a), b) and c) from the interpreted conversational behavior) and BEL elements from the beliefs-base, and apply *belief formation rules* and *belief updating rules* taken from the belief rules-base to these elements. As a result, new beliefs may be formed or existing beliefs may be altered.

In this section we specify in more detail how the rules from the various rules-bases function and what elements they need to form new elements or alter already existing elements.

Categorization rules: As mentioned in the description of the categorization rule-base, the *categorization rules* function by checking if properties a) to d) assigned to the conversational behavior during the interpretation match any of the properties that define the categories. One categorization rule exists for each of the properties. As mentioned earlier, the *categorization rules* only compare the properties of the behaviors with those of the categories for their level 1 values, as the level 2 values of the behavior properties are too specific to allow categorization.

When at least one of the category properties matches one of the `behavior properties` of the conversational behavior, that behavior can be placed into that category. Consequently, it is possible that a single conversational behavior can be grouped into more than one category, depending on its properties.

Belief formation rules: The formation of new beliefs about the situation that is expressed in B's conversational behaviors occurs in the belief formation and updating processes. Here, *belief formation rules* manipulate the level 1 and level 2 values of the properties a), b) and c) of the conversational behavior (contained in the categories in the behaviors-base) to construct new beliefs.

Some of the *belief formation rules* are used to convert the `intended effects` of the conversational behaviors performed by B (i.e. property a)) into new **beliefs** of A and store them in A's belief-base. The contents of these new beliefs cover the **intentions** that A ascribes to B based on the intended effects of B's conversational behaviors. For example, the intended effect of a conversational behavior (performed by B) grouped in the Gather information category is to *`get information`* from A. The intention that A associates with this behavior property, is that B wants A to perform a conversational behavior from the Inform about bad/good situation category. The *belief formation rules* convert the intended effect of the conversational behavior into the following belief: $BEL_A(INT_B(ToPerform_A(\text{Inform about bad/good situation})))$, 0.7. The variable x expresses the exact situation that B wants A to convey.

In addition to forming new beliefs about B's intentions, the *belief formation rules* also convert the `expected effects` (i.e. property b)) of the conversational behaviors performed by B into new beliefs. The expected effects are related by A to the expectations that A believes are present in B's internal state. As a result, the *belief formation rules* will produce beliefs about expectations: $BEL_A(BEL(exp)_B(x), 0.7), 0.7$.

The belief formation and updating processes in the dialogue model also use certain *belief formation rules* to convert the configuration of the internal state features that (according to A) underlie the selection of a conversational behavior, into new beliefs of A. This is done by simply replacing the variable x in $BEL_A(x)$, α with the internal state features that make up the configuration. For example, interpreting the conversational behavior "Did you take your medication?" performed by B results in A ascribing the following configuration of internal state features (that is, property c)) to B:

- BEL_B(A has taken his medication), 0.1.
- $BEL_B\neg$(A has taken his medication), 0.1.
- $BEL_B(BEL_A$(A has taken his medication), 1.0 or $BEL_A\ \neg$(A has taken his medication), 1.0), 0.7.
- $BEL(exp)_B$(ToPerform$_A$(Inform about good/bad situation((A has taken his medication) or $\neg$(A has taken his medication)))), 0.7.
- $DES_B(BEL_B$(A has taken his medication), 1.0 or $BEL_B\neg$(A has taken his medication), 1.0).
- INT_B(ToPerform$_A$(Inform about good/bad situation((A has taken his medication) or $\neg$(A has taken his medication)))), 0.7.

The first two **beliefs** express that B is uncertain (i.e. both beliefs have low certainty values) whether A has taken his medication or not. The third element expresses that B believes that A knows (i.e. believes with a certainty value of 1.0) whether A has taken his medication or not. The fourth element is an **expectation** that expresses that B expects A to perform a conversational behavior from the Inform about good/bad situation category as a response. Expressed by the fifth element, is B's **desire** to know (i.e. believes with a certainty value of 1.0) whether A has taken his medication or not. The last element in the configuration expresses that B aims (has the **intention**) to get A to give him the wanted information by performing a conversational behavior from the Inform about good/bad situation category.

These six elements are used as input for the belief formation processes, which use *belief formation rules* to convert these elements ascribed to B's internal state into the following new beliefs:

- $BEL_A(BEL_B$(A has taken his medication), 0.1), 0.6.
- $BEL_A(BEL_B\neg$(A has taken his medication), 0.1), 0.6.
- $BEL_A(BEL_B(BEL_A$(A has taken his medication), 1.0 or $BEL_A\neg$(disease is curable), 1.0), 0.7), 0.6.
- $BEL_A(BEL(exp)_B$(ToPerform$_A$(Inform about good/bad situation((A has taken his medication) or $\neg$(A has taken his medication)))), 0.7), 0.6.
- $BEL_A(DES_B(BEL_B$(A has taken his medication), 1.0 or $BEL_B\neg$(A has taken his medication), 1.0)), 0.6.

- $BEL_A(INT_B(ToPerform_A(Inform\ about\ good/bad\ situation((A\ has\ taken\ his\ medication)\ or\ \neg(A\ has\ taken\ his\ medication))))$, 0.7), 0.6.

The certainty values of A's newly formed **beliefs** are calculated by taking a base value (base) and adding the value of the **trust** A has in B (represented by τ) plus the assumed level of probability (p) that the content of the conversational behavior (i.e. x) is true. From this a normalized value about the number of beliefs of B, that have contradicted beliefs of A earlier in the conversation (represented as ϕ) is subtracted. The value ϕ is calculated by dividing the number of beliefs that contradict the existing beliefs by the total number of beliefs A has. The total summation of variable values is then divided by the amount of variable values (in this case four) to normalize the certainty value. This is expressed in the following formula:

- α = (base + τ + p ϕ) / 4 (Certainty value for beliefs formed based on the properties of the conversational behavior).

Before the start of the conversation, the base value of the certainty needs to be established. In our examples we have arbitrarily set the base value to 0.5, but this can be varied should that be necessary.

In addition to forming new beliefs based on the properties of the conversational behaviors performed by B, A can also use one of the *belief formation rules* to form new beliefs based on his existing ones. Often this rule is applied to existing beliefs that are about beliefs of B, i.e. $BEL_A(BEL_B(y), \beta), \alpha$ but only if the certainty value of A's existing belief (i.e. α) and the certainty value of the belief that A has ascribed to B (i.e. β) are both high. If this is the case, then A can adopt the belief he ascribed to B as his own belief. In other words, if A knows (= believes with a high certainty) that B knows (= believes with a high certainty) something, A can adopt this knowledge for himself.

The *belief formation rule* for A adopting beliefs he ascribed to B as his own beliefs operates as follows: The rules checks whether both the α value and the β value of the target element (i.e. $BEL_A(BEL_B(y), \beta), \alpha$) are both above the threshold of 0.8. Next, the attribute part (i.e. BEL_B) of the value that has been given to the x variable of A's belief (i.e. $BEL_B(y), \beta$) is removed along with its corresponding certainty value. What is left after doing so is A's new belief: $BEL_A(y), \alpha$.

Belief updating rules: Next to the *belief formation rules*, the belief formation and updating processes also operate by applying *belief updating rules* to elements from the belief-base. As mentioned earlier there are two kinds of *belief updating rules*. The first kind is used to check whether newly-formed beliefs contradict any of the other beliefs present in the belief-base. The belief updating processes access the list of contradicting beliefs and use one of the *belief updating rules* to see if the new belief is part of one or more of the contradicting belief-pairs. If a contradicting belief-pair is found in which the new belief matches one of the two parts, another *belief updating rule* is applied. This second *belief updating rule* checks if the other part of the contradicting belief-pair is currently present in the belief-base. If no match is found in either the list of contradicting beliefs or in the other beliefs in the belief-base, the new is placed in the belief-base.

If both parts of a contradicting belief-pair are found (i.e. the new belief and an existing belief), a third *belief updating rule* is applied to delete the belief with the lowest certainty value of the two from the belief-base. One exception to this rule is that if both beliefs have certainty values that are very low (representing that A is very unsure about what he believes), both beliefs are kept in the belief-base. Another exception is if the contradicting beliefs both have the same, high certainty value. In that case, both beliefs are also kept in the belief-base, but both their certainty values are halved. This represents an increase in the uncertainty of those beliefs.

A problem occurs when two beliefs, an old one and a new one, contradict each other but no contradicting belief-pair was constructed for them. In the current model, this means that the two beliefs are not recognized as contradictory. As a result both beliefs are kept in the belief-base, possibly causing inconsistencies in A's conversational behaviors. This is still an open issue in the dialogue model.

The second kind of *belief updating rules* is used to make small alterations to beliefs in the belief-base. These alterations include increasing or decreasing certainty values, or adjusting the intensity values in A's beliefs about B's emotions. For example, A believes that B is experiencing the emotion Sadness with intensity 0.3 (i.e. B is only a little sad) based on previously performed conversational behaviors. Then B performs a new conversational behavior which contains a strong display of Sadness. Consequently, A may alter his current belief about B's sadness, so that it indicates B is feeling more sad, by adjusting the intensity value of the emotion in his (i.e. A's) belief to say 0.6. When this *belief updating rule* is applied in this manner, the certainty value of the belief is not adjusted. By applying the same rule again the certainty value might also be altered.

Intention updating rules: Over the course of the conversation, A can decide that he needs to commit to another desire meaning that he needs to alter his intention. There are several situations which cause such an occurrence to happen.

One situation which causes A to alter his intention is when it becomes clear that the desired situation is unattainable. This is represented as an **expectation** with a low certainty value (as the probability p of the situation becoming true is low) in A's belief-base.

Another situation is that the desired situation is already brought about, either through conversational behaviors performed by A or through any other means. A situation that already has been brought about, i.e. holds true, is represented as a **belief** with a high certainty value in A's belief-base. After all, if A is certain in his belief that the situation is true, he will drop the desire to bring the situation about. This may even occur if the situation is not true, but A believes it to be so.

A third situation in which A can decide to alter his intention is when another desire is more appropriate or easier to obtain, given the current situation. A desire is more appropriate or easier to obtain if the conversational behavior that brings about the desired situation can be selected more easily. A conversational behavior can be selected more easily if more of its preconditions can be filled by elements from the internal state of A than for the conversational behavior that brings about the situation that fulfills the current desire. If the same amount of selection premises can be filled

for two or more conversational behaviors, one of the associated desires is selected at random to commit to.

The *intention updating rules* check the belief-base to see if the conditions of any of the situations hold. If they do, the *intention updating rules* will drop the current intention and select another desire to become A's intention. Dropping and committing to the more appropriate desire is simply a matter of changing the attribute of the relevant elements, either from DES_A to INT_A, or vice versa.

Emotion state update rules (appraisal rules): The process of emotion appraisal, i.e. the emotional assessment of a situation expressed in the semantic content of a conversational behavior, is performed by using *emotion appraisal rules*. The *emotion appraisal rules* evaluate how A's disposition relates to his interpretation of a situation expressed by a conversational behavior performed by B. Alternatively, the rules evaluate how A's disposition relates to his internal state after a change in A's internal state occurs that is not directly caused by an external stimulus such as a conversational behavior performed by B. For example, if the formation of a new belief causes A's current intention to be achieved, A might experience a feeling of Joy. The outcome of the evaluation determines if the intensity values of the relevant emotions in the emotion state-base need to be increased.

In order to evaluate the relation between A's disposition and the interpretation of a situation, the *emotion appraisal rules* make use of elements from the desire-base, the social relation-base of A and either property g) of an interpreted conversational behavior or A's belief-base in addition to his desire-base and the social relation-base. The relation between A's disposition and a situation that occurs, or his internal state, can be expressed in terms of appraisal variables (also known as emotion eliciting condition relations).

Whenever a conversational behavior is interpreted, the appraisal processes use the semantic content (i.e. property g)) and elements from the mentioned components to determine what the values of the appraisal variables are. However, because this task requires the interpretation of the semantic content of the conversational behavior and knowledge about how that content relates to A's disposition, there are no *emotion appraisal rules* included for this in the emotion state rules-base. This is because, although this task seems intuitively easy for humans, it is actually really hard to represent in a model. Instead the allocation of the appraisal variable values in the dialogue model needs to be done by humans. The outcome of this process is a particular configuration of appraisal variable values that are associated with the situation expressed by the conversational behavior.

Each type of emotion in the emotion state-base also has a fixed set of appraisal variables and values that is unique for that **emotion type**. The *emotion appraisal rules* function by checking if the values of the appraisal variables associated with the conversational behavior (which are assigned based on the interpretation of the semantic content) match the values of the appraisal variables determined for each emotion. If the values match, then that emotion is triggered by the situation. A particular situation can trigger zero, one or even more than one emotion at the same time.

The appraisal variables that are included in the dialogue model have been determined by comparing several theories and models from the literature and seeing which variables are similar. As mentioned, studied theories and models include the OCC model (Ortony et al., 1988), Clark Elliot's Affective Reasoner model (Elliot, 1992) and the Émile system (Gratch (2000), Marsella and Gratch (2006)). The complete set of appraisal variables contained in the dialogue model is the following:

Appraisal variables:

- ***Desirability-for-self:*** This variable indicates to which degree A's desires to see the situation or environment in a particular state. It is an assessment of the situation with respect to his desires. If the situation causes the desire to become more unattainable, the value of this variable is *'undesirable'*, whereas if the situation fulfills the desire or brings A closer to fulfilling it, the value of this variable is *'desirable'*.
- ***Desirability-for-others:*** This variable expresses A's interpretation of whether the situation is desirable for B. The value of this variable can be determined by looking at the desires assigned to B by A. These desires are represented in the description of property c) of the conversational behavior. By checking the desires underlying B's conversational behavior, A can determine whether the situation described in the behavior is either *'desirable'* or *'undesirable'* for B.
- ***Status:*** This variable indicates how the new situation (expressed through the conversational behavior performed by B) relates to A's expectations. It shows whether the conversational behavior *'confirms'* or *'disconfirms'* the expectations A had of the situation prior to the conversational behavior. Another possibility is that B's behavior does not say anything that is relevant to A's expectations. Then the value of the status variable is denoted as *'unconfirmed'*, leaving the possibility open for confirmation or disconfirmation later on.
- ***Likelihood:*** This variable expresses the *'probability value (p)'* of the situation being true.
- ***Praiseworthiness:*** This variable indicates whether A or B should be assigned an attribution emotion. This variable depends on who was responsible for the situation expressed in the conversational behavior (which in case of A could lead to Pride or Shame and in case of B to Praise or Blame to be experienced by A), but it may also involve considerations of intention. For example, if B unintentionally did something blameworthy, A will respond differently than when B did it intentionally.
- ***Appeal:*** This variable indicates the degree of A's liking or disliking towards objects (including people) in the situation. Its values are *'liking'* or *'disliking'*.

Only those appraisal variables that are appropriate given the situation expressed in the behavior are given values. For example, when B says "You are going to be fine." the focus of the situation is A's well-being. Subsequently, when A appraises this conversational behavior, the variables ***desirability-for-self*** and ***likelihood*** seem appropriate to get values and possibly ***status***, depending on the expectations A had formed prior to B's behavior. A variable like ***desirability-for-other*** does not apply to this particular conversational behavior nor does the variable ***praiseworthiness*** (unless

A believes that he is going to be fine as a result of B's actions) and thus they do not get assigned values.

Contained in A's internal state is the desire to be healthy, i.e. being in the situation where the state of his health is positive. This leads to the value of *'desirable'* to be assigned to the ***desirability-for-self*** appraisal variable, as the content of the behavior aims to fulfill that desire. Also, as the conversation is held between a doctor (B) and his patient (A), it is also reasonable to assume that B is knowledgeable and is telling the truth and therefore the ***likelihood*** of the state of the situation being true is quite high.

By comparing these values of the appraisal variables with the set values of the different emotion types, we find that A is experiencing the emotion Hope as a response to the conversational behavior. This is because the values of the appraisal variables are similar as the values contributed to the Hope emotion type. This is expressed by an increased intensity value of the element in the emotion state-base that represents the emotion **Hope**.

In addition to the *emotion appraisal rules* that may cause an increase in the intensity values of the emotion elements, the emotion state rules-base also contains a decaying function that causes the intensity values to decrease steadily over time. This decaying function represents A's habituation to a particular situation and subsequently his lessening emotional response to it.

Social relation updating rules: The social elements that together represent the way A perceives his interlocutor B socially (i.e. his social relation or social disposition with respect to B) can be affected by applying *social relation updating rules* to the information offered to the social relation updating processes. The way in which a particular *social relation updating rule* functions depends on the source of the provided information.

The information provided by the belief-base indicates how many of the newly formed beliefs, that are based on B's conversational behaviors, contradict with the beliefs A already possessed. This information has a strong influence on the intensity value of the social aspect type **trust**. If the new beliefs that A forms based on the processing of B's conversational behaviors contradict with the beliefs A already had in his belief-base and A retains his old beliefs instead of replacing them with these newly-formed beliefs, A's trust in B will decrease. A will retain his old beliefs if they have higher certainty values than the newly-formed beliefs.

Information from the emotion-base can be used to influence the intensity values of the **liking** and **rapport** social aspect types. For example, if A is experiencing a negative emotion such as Anger or Disgust with respect to B, this will cause A's **liking** of B to become less. After expressing the emotion, the related *social relation updating rule* will also decrease the intensity value of the **rapport** that exists between A and B as perceived by A, as this will most likely be a consequence of expressing the negative emotion.

Property g) of the conversational behavior (i.e. the semantic content) will provide information about the degree of politeness and the degree of formality of the conversational behavior based on the words used. Other displays of B's social dispo-

sition with respect to A, such as proximics and gestures, also determine whether the behavior is polite or not. As a result, the *social relation updating rules* will also take into account these other social displays, but instead of using the interpretation that ascribes a social disposition to B (i.e. property f)), A only checks how they express the degrees of politeness and formality.

The degrees of politeness and formality are not only determined by the language used (i.e. the selection of certain words) or by certain gestures, but also by whether the interlocutor adheres to social norms concerning the make up of a conversation. For example, the use of salutations, self-introduction and valedictions are examples of a polite conversation. Furthermore, interpretation of the degrees of politeness and formality can be influenced by the social roles both interlocutors have during the conversation. For example, people are inclined to be more polite when they hold their interlocutor in high regard socially, such as a tutor or a doctor.

The *social relation updating rules* check whether A perceives B's conversational behaviors as being polite and/or formal on the basis of a collection of set standards (e.g. is B using polite language? How friendly is the tone of the conversation? Etc.). Based on the evaluation of the conversational behavior with respect to politeness and formality, the *social relation updating rules* will increase or decrease the intensity values of **liking** and **rapport**, represented by λ and ρ.

Dialogue management rules: The process of updating the elements in the dialogue state base is done by applying *dialogue management rules* obtained from the dialogue management rules-base. Some of the dialogue management rules determine based on the `behavior type` of the interpreted conversation behavior, what possible behavior types the response behavior can have. Alternatively, such rules check if the behavior type of the conversational behavior was appropriate given the behavior type of previous conversational behavior.

The relation between two behavior types expressed in the *dialogue management rules* is based on the notion of adjacency pairs. Each of these rules consists of two behavior types and an operator that indicates how they are related to each other. For the sake of transparency, each operator only indicates a single, directional relation between the behavior types, indicating that either one follows the other or that one is preceded by the other. This approach facilitates the dialogue state updating processes, which only have to compare the first behavior type in the rule with the behavior type ascribed to the interpreted conversational behavior.

However, this also means that each pair of behavior types has two rules representing their relations. For example, it is possible that an *answer* is preceded by a *question* and followed by a *statement*. At the same time, the *question* is followed by an *answer* which precedes a *statement*. These example relations are represented in the dialogue management rules-base as follows:

- *BEH(answer)*, $\leftarrow$*(BEH(question))*
- *BEH(answer)*, $\rightarrow$*(BEH(statement))*

- *BEH(question)*, $\rightarrow$*(BEH(answer))*
- *BEH(statement)*, $\leftarrow$*(BEH(answer))*

When the behavior type of the conversational behavior matches the first `behavior type` of one (or more) of the *dialogue management rules*, it is simply a case of interpreting the operator and the second `behavior type` to see how the behavior type is related. If the rule contains an 'is followed by' operator, the relevant dialogue management rule will assign the *ToPerform* attribute to the second behavior type and store it in the dialogue state-base. If the rule contains a 'is preceded by' operator, the second behavior type is compared with the last behavior type added to the sequenced list in the dialogue state-base. If the two behavior types match, the first behavior is given the *ToPerform* attribute and added to the sequenced list. If the behavior types do not match, this indicates a discrepancy in the conversation. In such cases, a repair mechanism needs to be used to get the conversation back on the right track. The behavior type of conversational behavior is not stored and any suggested behavior types that have received the *ToPerform* attribute are overruled.

In addition to keeping track of the behavior types of performed and to-be-performed conversational behaviors, the dialogue state updating processes also keep track of the phases of the dialogue, i.e. start-middle-end. This guaranties that `behavior types` that are only appropriate in a certain dialogue phase are not used in other phases. Some of the *dialogue management rules* filter out suggested behavior types that are inappropriate. For example, performing a conversational behavior of the `behavior type` *greeting* in middle or end phases of the conversation should not occur, thus when such a behavior type is suggested, it is immediately filtered out again.

Behavior selection rules: As mentioned in section [6.2.1] the *behavior selection rules* are similar to the *categorization rules* only they operate in a reversed manner. In order to select an appropriate response behavior, the dialogue model tries to match the configuration of the internal state of A, consisting of his **beliefs**, **desires**, **intention**, **emotion state** and his **social state** to the (level 1) values of the properties that are ascribed to each of the categories in the behaviors-base. The more internal state features match the preconditions (i.e. the values of the category properties), the higher the chance is that a conversational behavior from that category will be selected as a response.

When a match is found between the internal state features and the property values of a category, the *behavior selection rules* select the conversational behaviors from that category for which the level 2 values are most similar to the level 2 values of the internal state features.

The behavior selection processes continue to narrow down the number of conversational behaviors by evaluating the information stored in the dialogue state-base (see the explanation of the *dialogue management rules* in this section for more information). The same narrowing occurs based on the intensity values of the elements in the emotion state. If an emotion is experienced with a very high intensity, it will override the previous selection of conversational behaviors in favor of a conversational behavior that is intended to express the strongly experienced emotion. As a result of these eliminations, the most appropriate response behavior (i.e. the one that is most in line with A's internal state features and with the dialogue state) will be selected and performed. The influence of the emotion state and / or the social state on the

form of the selected conversational behavior is applied in a later stage. However, this is not included in the dialogue model as realization of the behavior falls outside the scope of this model.

6.3 Related work

In the previous sections, we have presented the structure of our cognitive dialogue model and clarified the elements and rules that respectively represent features of the internal state and the cognitive processes that form or adjust them. As mentioned in the first couple of paragraphs of this chapter, the model has been created based on the study of various psychological theories, models, mechanisms and real life conversations in order to construct a model that was as realistic and cognitive plausible as possible.

The core components of the dialogue model (i.e. the belief-base and the desire-base) are based on Michael Bratman's BDI model of human practical reasoning (Bratman, 1987). As a result, the dialogue model as a whole is structured as an information state model. The BDI model was selected because it provides us with a structured basis which represents certain features of a person's internal mental state. The internal state features expressed in the BDI-model represent the rational part of a person's internal state, i.e. his knowledge about the state of the world / the situation and his aims to alter aspects of the state of the world in order to bring about a new one. In our dialogue model, these internal state features (i.e. the knowledge and aims) are represented in the same way as in the BDI model, i.e. as **Beliefs**, **Desires** and **Intentions**. We adopted the same notations for the elements in the dialogue model to indicate the internal state features expressed in the BDI model, respectively BEL, DES and INT.

However, the BDI model is somewhat limited, as it does not represent all aspects of a person's internal state. For example, it fails to take into account the influences emotions and social dispositions have on a person's reasoning about the states of the world. In order to construct a more human-like cognitive dialogue model, we argue that such internal state features need to be included in the representation of a person's internal state. Luckily, in the last couple of years there has been an increase in studies that try to integrate emotions into the BDI model. For examples see (Steunebrink, 2010), (Adam et al., 2006) or (Pereira et al., 2008). One of the most commonly used emotion models for extending the BDI model is the OCC model (Ortony et al., 1988), because of the clear classification of emotions it provides. The OCC model also contains BDI-like concepts to describe the conditions that elicit emotions.

As mentioned in section [6.2.2], we adopt the classification of emotions as described by the OCC model to specify the emotion elements that make up the emotion state in our cognitive dialogue model. In addition to the **22 emotion types** described by the OCC model, we have included the following internal state features as emotion elements in the emotion state of our cognitive dialogue model: **Surprise**, **Confusion**, **Sympathy** and **Empathy**. These internal state features are not included in the classification of the OCC model, but we argue that, rather than seeing them as a person's dispositions with respect to a certain situation, they can be classified as emotion

types. These four internal state features were found to be abundantly present in the analyzed bad news conversation.

Instead of adopting the three main categories from the OCC model (i.e. the consequences of events, the actions of agents or the aspects of objects), we make a distinction between two kinds of **emotion types** and divide them into two groups within the emotion state-base. The first group contains emotions that relate to the emotional disposition of the person with respect to himself. Such emotions may either be caused by a behavior or event in the world or a specific configuration of A's internal state features, but they are only directed at the person himself. Examples of such emotions include Joy, Sadness, Confusion and Surprise. Emotions contained in the second group are related to the emotional disposition a person (A) has with respect to his interlocutor (i.e. B). Opposed to the emotions in the first group, these emotions are actively directed at something or someone other than A. This holds for emotions such as Anger, Sympathy, Praise and Blame.

Although we use the specification of emotions of the OCC model to indicate which **emotions** are included in the dialogue model, the variables used in the appraisal processes of our dialogue model are constructed based on the variables of multiple emotion theories and models. Prior to the construction of the emotion appraisal processes, we looked at the appraisal variables of several emotion models that are used in various conversational agent systems. In addition to the OCC model, we also studied the Affective Reasoner process model (Elliot, 1992), the Émile architecture (Gratch, 2000) and the EMA process model (Marsella and Gratch (2006), Marsella and Gratch (2009)) (which builds further on the Émile architecture), to determine which appraisal variables would be most appropriate to use in our cognitive dialogue model. As a result, we found that each model contained several appraisal variables that were similar to appraisal variables in one or more of the other models, if not in exact description than in meaning. We argue that these similarities indicate a consensus about which variables are indispensable for the process of appraisal. Consequently, we only used the appraisal variables that were presented in multiple models to be part of the appraisal processes in our dialogue model.

As shown in chapter [4], the values of the appraisal variables are assigned after the situations in the conversational behaviors are interpreted. The appraisal processes of the studied models are explained by means of discrete situations occurring in clear and simple domains. Because the example situations are so straightforward, the assignment of values to the appraisal variables is taken for granted.

However, studying the models did not provide us with an explanation how the processes of interpreting the situations, needed for assigning the proper values to the appraisal variables, actually operate in the studied models. Furthermore, as our cognitive dialogue model aims to represent the natural, and often ambiguous, conversational behaviors in the complex domain of bad news conversations, interpretation of the situations described in the conversational behaviors is considered to be more difficult. Due to this increased difficulty, the lack of comparable representations in other models and the constraint of time, the process of interpretation is not included in our dialogue model. Interpretation of the situations is done by human observers. As a result, the values of the appraisal variables are also assigned manually.

The evaluation procedure adopted by the emotional appraisal processes in our model is similar to those used in the studied models. The processes of appraisal operate by comparing the configuration of the values assigned as a result of the interpretation with fixed configurations of values that are associated with each individual emotion type. For our model we adopted the fixed configurations of values (for the previously selected appraisal variables) for each of the 22 emotion types described in the OCC model. The fixed configurations of values for the additional four emotion types we included in the model, were constructed based on the other configurations and common sense reasoning.

In addition to the internal state features that represent the emotional disposition of a person, the dialogue model also includes representations of social aspect types that together describe the model's social disposition with respect to its interlocutor. As mentioned in section [6.2.1], the selection of these social internal state features was for a large part inspired by work by Falcone and Castelfranchi (2001) and Bickmore (2003) (for the social feature **trust**), and Gratch et al. (2006a) (for the social feature **rapport**). Although **liking** is a social aspect that is often used in studies to indicate how much a person favorably regards the conversational system or intelligent agent he is interacting with, we argue that this social aspect can also be used to represent a part of the social disposition present in the dialogue model. As a result, we included the social feature **liking** in our dialogue model.

Instead of adopting the representations used in the mentioned studies to express the social aspect types, we constructed elements that are more in line with the other representations of internal state features to represent the social internal state features. This was done in order to ensure compatibility between the different elements and to enable the cognitive processes to combine the various types of elements. Assigning values to these social features was done manually based on the interpretation of the conversational behavior with respect to the model's disposition towards the interlocutor.

After performing a thorough study of the literature we came to the conclusion that few dialogue models clarify in detail the processes involved in the selection of appropriate conversational response behaviors. Nor do they offer clear representations of complex conversational behaviors and the internal state features associated with them, that are often performed in realistic, natural dialogues. Although these shortcomings have provided us with the motivation to construct our cognitive dialogue model in the way it is presented in this thesis, they also make it difficult to compare our dialogue model with other dialogue models in order to evaluate it.

The performed literature study indicated that the structure of our cognitive dialogue model resembles architectures of intelligent agents more closely than it does other dialogue models. Hence it seems a better idea to make a comparison between our model and such structures. We found that both the agent architecture of the EMA system (Marsella and Gratch (2006), Marsella and Gratch (2009)) and the MicroPsi architecture (Bach, 2003) showed several similarities to our dialogue model. As can be seen in figure [6.3], the architecture underlying the EMA system contains a collection of internal state features that is based on the interpretation of the relationship between the agent and the environment, called the *causal interpretation*. Similar to

our own cognitive model, the internal state in the EMA architecture consists of **Beliefs**, **Goals** (or **Desires**) and **Intentions** and is augmented with the features Causal Relations and Plans. This causal interpretation controls the selection of the actions and dialogue behaviors of the agent. The selected actions and dialogue behaviors are subsequently moderated by control signals, which are the result of the process of emotion appraisal. As mentioned earlier in this section we adopted some of the appraisal variables presented in the EMA model to be used in our own emotion appraisal processes.

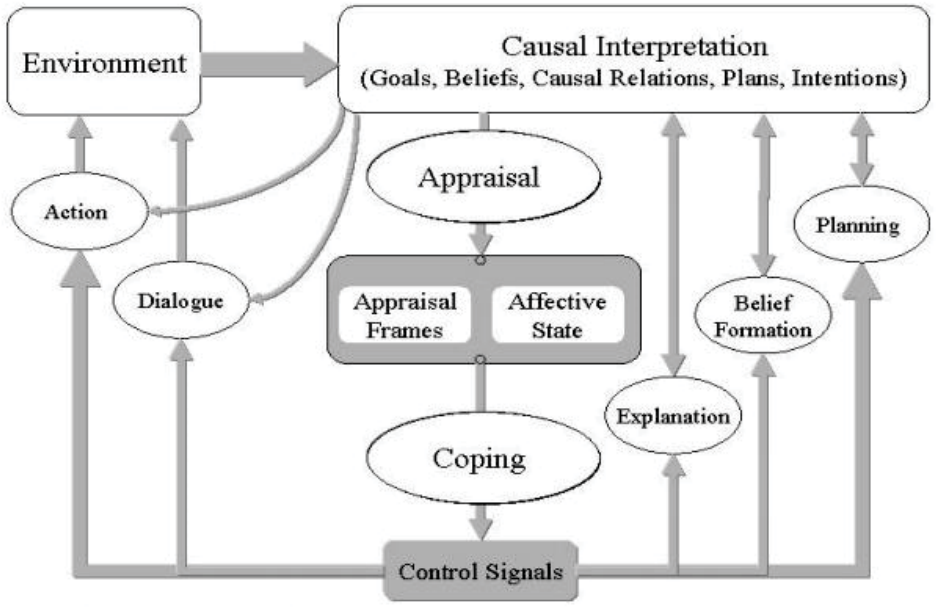

Figure 6.3: The EMA architecture.

Our cognitive dialogue model functions differently from the EMA architecture in several ways. For one, our dialogue model allows for the formation and updating of the BDI internal state features (i.e. **Beliefs**, **Desires** and **Intentions**) without the influence of control signals that are formed by the appraisal and coping processes. Furthermore, the **emotions** in the affective state in our dialogue model have a direct influence on the process of behavior selection, rather than playing a modifying role which is performed through the control signals. In addition, we also take into account how the social disposition of the model's agent influences the selection of appropriate response behaviors.

As can be seen in figure [6.4], the structure of the MicroPsi architecture (Bach, 2003) also contains features that are comparable to those present in our own cognitive dialogue model.

The MicroPsi architecture makes a distinction between two types of input to which the agent can respond. The concept of Urges is comparable to our **Desire** features. Urges leads, via the representation of Motivations, to the selection of appropriate **intentions** and the selection of behavior. Instead of dealing with complex dialogue situations or cognitive states of the interlocutor that the agent wants to bring about,

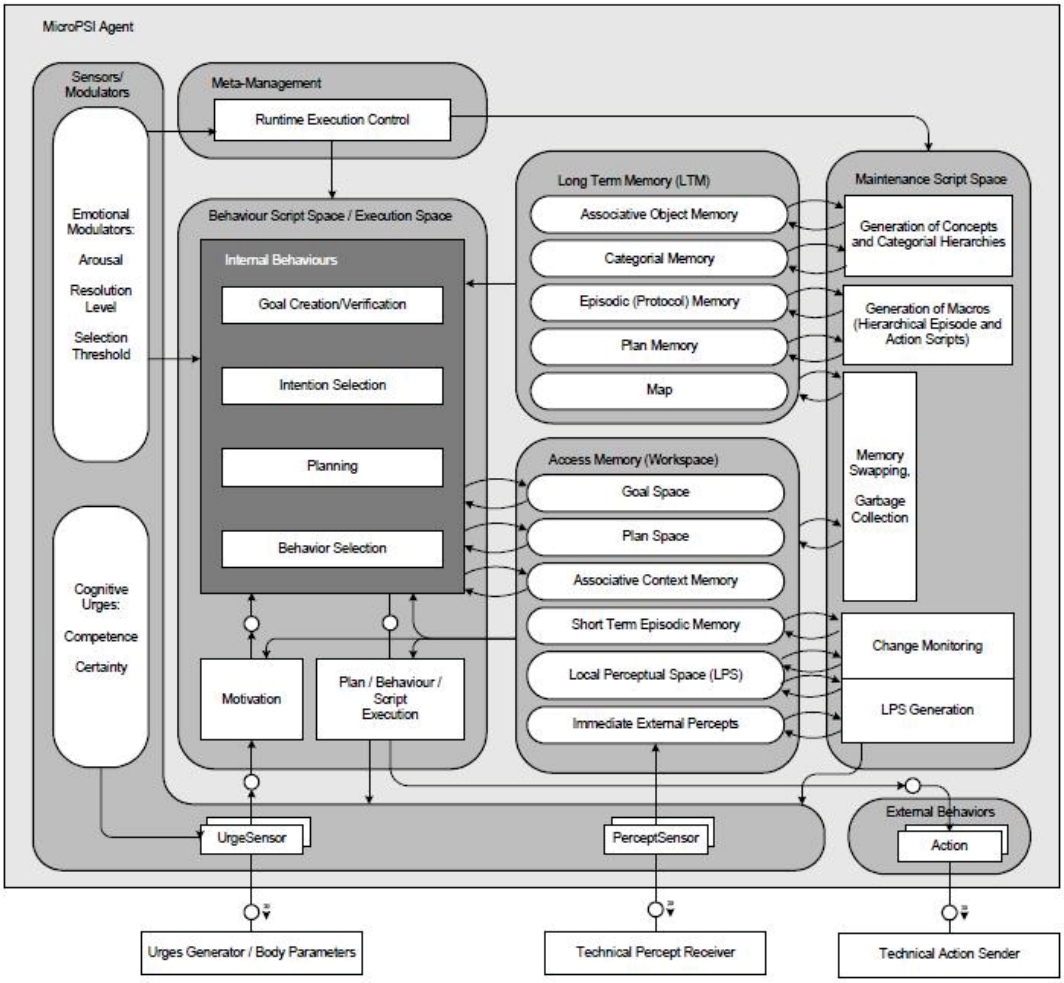

Figure 6.4: The MicroPsi architecture.

the urges in the MicroPsi system are concerned basic needs such as energy (food and water), intactness (i.e. safety of the agent) and competence. In addition, they include more cognitive notions such as affiliation (i.e. social satisfaction through inclusion) and the reduction of uncertainty. These latter concepts also are present in our own dialogue model. Affiliation as the social aspect types **rapport** and **liking**, and the reduction of uncertainty indirectly through the selection and performance of conversational behaviors with the `intended effect` *`question`*.

The other type of input (i.e. percepts) represents the current state of the agent's environment and manipulations that occur therein which might be caused by either the agent itself or for example by an interlocutor. Processing and storing of the representations of the environment occurs in the Access Memory (Workspace) in the Local Perceptual Space and the Short Term Episodic Memory components. These components can be seen as alternatives for the Belief-base we have included in our dialogue model and their representations as the **beliefs** internal state features. Through further processing the MicroPsi architecture can convert the representations of its knowledge into features stored in the different types of memory components of the Long

Term Memory (LTM), such as episodic memory, categorical memory, plan memory and associative object memory. In our dialogue model, the only distinction we make is between **beliefs** about events and states that already have occurred and **beliefs (expectations)** about events and states that may come about in the future.

The Goal Space and the Plan Space components in the Access Memory (Workspace) have a similar function to that of the Desire-base component of our dialogue model, namely storing the agent's possible **intentions** and **desires** (i.e. goals). The content of the Goal Space and Plan Space components is used by the Goal Creation and Intention Selection processes. In addition, the creation of goals and the selection of intentions in the MicroPsi architecture is based on the motivations rather than the desires (i.e. Urges) of the agent. The Goal Creation and Intention selection processes correspond with our Intention updating processes. Our dialogue model lacks the capabilities of creating new goals, but is restricted to selecting a single one from a predetermined list to become its current intention.

The emotion modeling in the MicroPsi architecture is based on a framework created by Dietrich Drner called the Psi-theory. (For more information about Drner's Psi theory see Bach (2003)). The framework operates under the assumption that the emotion state of the agent is expressed as a set of configurations of the cognitive system. The configurations are defined by certain operators called modulators, such as arousal, resolution level and selection threshold. Whenever a change in the modulators occurs, a particular configuration of the internal state of the system influences the way an agent perceives, plans, memorizes, selects intentions, acts etc. In our own dialogue model such modulators are expressed as various internal state processing rules that use the internal state features contained in the emotion state-base (i.e. the **emotion types** and their intensity value) as premise. For example, the selection of an **intention** can be influenced if an emotion type has a particular high intensity value. In such a case, the selection processes will call upon a different set of rules than they would normally do. The same holds for the processes involved in the selection of appropriate response behaviors. At the same time, the features in the emotion state-base correspond to the set of configurations of the system mentioned in the Psi-theory. The advantage of using various emotion types instead of a set of configurations is that the emotion types are easier to comprehend for someone who uses the dialogue model, as they are more intuitive. In addition, the emotion types are easier to define clearly as they are individually represented and do not depend on underlying relations or processes.

One of shortcomings of the MicroPsi architecture is its limited ability to deal with the social aspects of conversational behavior. Because the architecture does not have the abilities to derive and assess features of the mental state of others, it is unable to represent and process any social relations that may exist between itself and its interlocutor. Since we included the capacity to take into account the internal state features of the interlocutor during the processing of perceived conversational behaviors and reasoning about how the performance of a specific response behavior might affect the internal state of the interlocutor, our cognitive dialogue model does allow for the representation of any social relations that may exist. We believe that the inclusion of social aspect types (**rapport**, **liking** and **trust**) and the possible influence they have

on the selection of appropriate response behaviors allows for a dialogue agent that is constructed based on the dialogue model to perform in a more natural and humanlike manner.

Although the comparison between the structure and features of our cognitive dialogue model and the architectures of intelligent agents indicate some of the advantages and disadvantages of our model, it is important to remember that a clear understanding of the form and function of a conversational behavior is also necessary in order to properly process that behavior. Not only comprehension of the content and the purpose (or function) of a conversational behavior is important, but also an understanding of the internal state features of the speaker that give rise to the conversational behavior. Such characteristic can often be determined by interpreting and processing the properties of the perceived behavior. However, enabling the dialogue model to interpret and process the properties of every conversational behavior requires a large amount of individual rules. It seems more practical to group behaviors that have similar properties together.

Consequently, we created a set of categories into which perceived conversational behaviors can be grouped. Each category is defined by a set of representations of behavior properties that hold for each conversational behavior contained in that category.

In order to determine which sets of behavior properties should be used to define the different categories, we studied the distribution expressed in the DIT++ taxonomy (Bunt, 2009). However, the elements in the DIT++ taxonomy are limited to describing just the communicative functions of possible conversational behaviors. As shown earlier in this chapter, a conversational behavior can provide the dialogue model with much more information than just its communicative function. In section [6.1.1], we have explained that in addition to the intended effect (i.e. the communicative function), three other types of behavior properties also play an important role in deriving the purpose of a conversational behavior and the mental disposition that underlies it. The following `behavior property` types are used to establish a proper understanding of conversational behaviors:

- The expected effects of the conversational behavior as anticipated by the speaker. The expected effects property indicates what expectation-beliefs hold in the speaker's internal state.
- The configuration of internal state features that need to hold before the speaker will select and perform the conversational behavior, such as his beliefs and desires.
- The behavior type of the conversational behavior.

These `behavior properties` are not taken into account in the elements presented in the DIT++ taxonomy. Hence, we argue that the DIT++ taxonomy is insufficient to be used to adequately process a perceived and interpreted complex conversational behavior. Instead, we made a selection of communicative functions from the DIT++ taxonomy that seemed most likely to be used in a bad news conversation and extended those elements to include representations of the other three behavior property types. The extended elements were then defined as the categories that make up the

categorization. The DIT++ taxonomy also does not include communicative functions that represent a person's intentions to share one of his emotions. Because of this, we needed to construct additional categories that could contain conversational behaviors that expressed such intentions through their intended effects property.

By being able to describe the purpose and the provenance of a conversational behavior in this extended manner our cognitive dialogue model can provide a more comprehensive image of the interlocutor's internal state features involved with his conversational behavior. Furthermore, it allows the dialogue model to select conversational response behaviors that are more appropriate given its own configuration of internal state features than without this information.

6.4 Implementation

In an attempt to make a beginning of this complete dialogue system, a simple implementation was constructed. To achieve this we used Flipper (ter Maat and Heylen, 2011). Flipper is a specification language and interpreter for Information State Update rules that can be used for developing spoken dialogue systems and embodied conversational agents. The Flipper system uses XML-templates to represent an information state (such as the dialogue model's internal state) and the rules that modify elements in that information state.

As such, the Flipper system is particularly suited for the implementation of cognitive internal state. The same holds for the representation of the cognitive processes in the dialogue model, which can be expressed through Flipper's types of XML-templates.

Each XML-template consists of a set of preconditions that need to hold (i.e. be true) before the template can be executed. To see if the preconditions are true, the system compares them with elements that are present in the information state. Checking is done most often by using comparison operators (=, <, >, etc.), possibly with the appliance of one or more simple arithmetic operations to combine the numerical values of certain elements.

In addition to the preconditions, each XML-template contains a list of effects that are brought about when the XML-template is executed. Possible effects that may occur are modifications made to the information state. This includes adding, removing or altering variables (for example **beliefs** or **desires**) or their values (e.g. updating a certainty value of a belief or a intensity value of an emotion), or the selection of a conversational behavior that can subsequently be executed by a realizer component of a dialogue system such as a conversational agent. Below we present an example of an XML-template that represents a *belief formation rule*:

```
<template id="BF1_1" name="Belief formation rule 1_1">
  <preconditions>
    <compare value1="$tag.D1_1.desire.Other.wellbeing_is_fine.name"
      value2="wellbeing_is_fine" />
  </preconditions>
  <effects>
    <update name="belief.new.desire.Other.wellbeing_is_fine.name"
```

```
      value="wellbeing_is_fine" />
    <update name="belief.new.desire.Other.wellbeing_is_fine.name.value"
      value="0.5 + $current.social_state.rapport.value +
      $current.social_state.liking.value + $current.social_state.trust.value" />
  </effects>
</template >
```

The representation of the internal state elements in the templates differs slightly from the representations used in the dialogue model. This is done to increase the readability of the templates. One of the major differences is the use of Self instead of the patient and Other instead of the doctor, when indicating to which interlocutor the element belongs. This allows for the implementation to also be used in conversations in other domains as well.

The attribute "tag" is used to indicate that the precondition is a internal state feature ascribed to Other's (i.e. the doctor's) based on Self's (i.e. the patient's) interpretation of Other's conversational behavior (most likely through the dialogue model's category property c)). In this case it describes Other's internal state feature **desire** 'to be fine'. If this precondition is considered *true*, then Flipper will execute the effects of the template. The two effects of the template indicate 1) the formation of a new **belief** about Other's **desire** and 2) establishing the certainty value of the belief based on the values of the social state aspects, **rapport**, **liking** and **trust**.

Up till now, the small implementation effort has produced several dozen rules that either form new **beliefs** and adds them to the internal state, or alter **beliefs** that are already present in the internal state. Also 26 rules that update the intensity values for each individual **emotion** in the internal state have also been included. Based on these preliminary attempts, we conclude that our cognitive dialogue model functions as planned and that implementation is viable.

6.5 Conclusions

In this chapter we described the categorization of conversational behaviors and the cognitive dialogue model that have been constructed in order to gain more insight in the relationships between the properties of a conversational behavior, the internal state features that underlie and have given rise to the behavior, and the internal state features that subsequently can be formed or updated. In addition, we aimed to contribute to the field of dialogue processing by cognitive intelligent agents, by making clear representations of the processes and features related to this topic.

In section [6.1], we argued that by using the categorization of conversational behaviors, a model or architecture does not need to possess a huge amount of *behavior processing rules*. Often, such *behavior processing rules* are configured to be able to handle the many different conversational behaviors that may occur in a natural conversation. Instead, the make-up of the categories presented allows for different conversational behaviors to be processed in the same way based on the similarities in the value of their `behavior properties`. This allows for a vast reduction of *behavior processing rules* as different processes can use the same type of input, namely

the categories, instead of each individual `behavior property`. In addition, the established relations between different internal state features associated with a particular conversational behavior remain available by grouping a conversational behavior into a particular category. In cases where the configuration of the values of the `behavior properties` does not fit an existing behavior category, it is quite easy to construct a new (or adapt an existing) behavior category to accommodate the new conversational behavior.

We conclude that processing multiple characteristics of a conversational behavior (e.g. its `intended effect`, its `expected effect` and the internal state features associated with it) is easier and more efficient when a categorization is used than using a complete set of rules for each new conversational behavior.

The cognitive dialogue model presented in section [6.2], is the result of an accumulation of information obtained from the study of various theories that focus on the processes involved in the processing and generating of conversational behaviors, augmented by additional insight that was gained from our own analyses of (bad news) conversations concerning the way people respond conversationally in such situations.

In section [6.3], we identify the points on which our cognitive dialogue model differs from the techniques and dialogue models that we have studied and make a comparison. One conclusion we draw from the study of the various techniques and methods, is that many of the internal state features (and the processes that handle them) are interconnected to each other and that such information needs to be incorporated in a dialogue model. However, many of the existing techniques, methods and dialogue models only focus on a specific part of processing the behaviors performed in a conversation, instead of incorporating this relational information. We argue that by focusing on a small part, one loses a great deal of information about the conversational behavior that is contained in the relations. As a response to this, we incorporated many of the processes and internal state feature types in our dialogue model and in section [6.2] we tried to clarify how they are interconnected with each other and how they might influence each other.

Another conclusion we draw from our study is that, in order for a dialogue model to make a proper assessment of the intentions underlying the interlocutor's conversational behaviors, it needs to be able to take into account the possible features of the interlocutor's internal state. We enabled our dialogue model to do so, by giving it the capability to form **beliefs** about the interlocutor's internal state features that may have led to the selection of the conversational behavior it has performed.

We also found that most of the studied dialogue models lack adequate representations of the social relations that hold between the two interlocutors during the conversation. We argue that these social relations are essential as they might influence many of the formation and updating processes occurring during the conversation. In our opinion this aspect of a conversation needs to be included in a dialogue model in order to make its processes cognitively plausible and its conversational behaviors more natural. Consequently, we represented three social aspect types in our dialogue model, namely **trust**, **liking** and **rapport**.

In addition we found that, because of the complexity of the interdependencies between the different types of internal state features and the various cognitive pro-

cesses, it is important that the *processing rules* and the internal state features are represented in a structured manner in the dialogue model. Otherwise, the compatibility between them may fail and the dialogue model will not be able to select appropriate response behaviors. Extensive evaluation and readjustment of the dialogue model's components is necessary in order to ensure that they are suitable for processing all the perceived conversational behaviors.

Furthermore, in order for the dialogue model to perform adequately, the dialogue system that uses it as its processing architecture needs to be able to properly perceive and express the various internal state features that underlie the conversational behavior. This goes a long way to reinforcing the notion that the dialogue model underlying the dialogue system is indeed cognitively capable of holding a (bad news) conversation in a natural and realistic manner.

7

Chapter 7: Utilizing the dialogue model

The dialogue model presented in the previous chapter offers a structured and systematic overview of the cognitive procedures performed by a person (or dialogue system) in order to process conversational behaviors. The conversational behaviors are characterized by distinct properties that offer a wealth of information about the thoughts and feelings of the speaker. In order to represent and classify these behavior properties, a taxonomy was constructed (see chapter [6]). The observed and classified properties of the interlocutor's behavior properties are converted into assumptions about internal state features that are associated with the properties via a categorization process. This categorization process provides the external input for the dialogue model.

As explained in chapter [6], the cognitive dialogue model contains representations of the internal state features (i.e. the elements in the model) and representations of the cognitive processes (i.e. the rules in the model). To clarify how the processes in the dialogue model operate, a step-by-step example of the workings of the dialogue model is provided in section [7.1], including the properties and elements that play a role.

In section [7.2], we present a number of concluding remarks concerning the workings of the dialogue model and about the presented example.

7.1 Working example of the dialogue model

In order to explain how the cognitive dialogue model works in practice, the processes involved in handling and performing conversational behaviors are written out in detail. Although the various processes in the dialogue model normally occur continuously, we describe them here in a sequential manner to make the explanation easier to understand. Each iteration of the process-cycle consists of three main steps, each of which can be divided into multiple sub-steps. During a conversation the dialogue model goes through several iterations of the process-cycle. The cognitive processes take place in the following steps (and sub-steps):

1. Observation and interpretation of conversational behavior performed by an in-

terlocutor.

1.1. Assigning behavior property values
1.2. Categorization of the conversational behavior

2. Forming / updating elements that represent internal state features in the dialogue model.

2.1. Updating the Dialogue state-base
2.2. Updating the Belief-base
2.3. Updating the Desire-base
2.4. Updating the Emotion state-base
2.5. Updating the Social relation-base

3. Selecting and performing appropriate response behaviors.

In this example, we only focus on the conversational behaviors from a segment of the bad news conversation held in the Pallium video, as an elaboration of the entire conversation would be unnecessary long. Furthermore, the work-out of the segment will provide sufficient insight in the exact workings of the dialogue model. In the example, only the features and processes of the patient's internal state are illustrated, as we envision that the cognitive dialogue model will be integrated to take on that role within a dialogue system that allows doctors to practice bad news conversations.

Instead of using the variables A and B to indicate the interlocutors, this example uses the variables P and D to indicate the patient and the doctor respectively. The P and D variables are also used in the representations of the elements and rules.

The interpretation of the behavior properties (performed by both the patient and the doctor) and the categorization of the conversational behaviors is based on the two analyses discussed in chapter [5].

Presented below is the transcript of all the speech utterances performed in the conversational behaviors. The dialogue turns from which the conversational behaviors will be processed in this step by step example are highlighted in the transcript.

Doctor (D1): *I'm about to visit John Filpot. He is recovering from surgery. He had a bowel tumor removed and evidence of metastases. This means that John's cancer is incurable. Now, I know the surgeon has talked with John, but I don't know if John really understands his condition.*

Patient (P1): *I'm feeling pretty good. Yeah I got some pain, but the surgeon got my tumor. It was attached to my bowel. But the most important thing is that the tumor is gone. I'll probably be out of here in a few days, once they can remove these tubes and get me on some decent food. I'm sure everything is going to be okay. No biggy, as my son says.*

D2: *Hey John.*

P2: *Hello doctor Rupert.*

D3: *I'm sorry to hear the news. We were hoping that surgery and a course of chemo would cure this tumor.*

P3: *What do you mean?*

D4: *When a tumor has spread through the abdomen, like yours has, there is a limited amount we can do for you.*

P4: *What do you mean, spread through the abdomen like mine has? The surgeon said he had removed the tumor.*
D5: *Yes, The primary one. But he couldn't get all the places it had invaded; (camera shift) we need to give you chemo for that.*
P5: *What do you mean invaded? You keep talking about this chemo-stuff.*
D6: *Didn't the surgeon talk with you?*
P6: *Yeah but he didn't say anything about chemo or invasive stuff. What the hell is going on here? I thought I was cured, (camera shift) but now it sounds like I'm not.*
D7: *I'm sorry John I thought you knew.*
P7: *Well I didn't! I don't believe it. (Pause) They have shown you the wrong chart! I'm going to be fine (camera shift). Just need a little medicine and I'll be right as rain.*
D8: *Okay, John. I'll be back later.*

Step 1) Processing the conversational behavior in dialogue turn D2:
Hey John.

As the first two dialogue turns are aimed at determining the background of the bad news situation, we argue that the actual conversation begins in dialogue turn D2. The conversational behavior in this dialogue turn is performed by the doctor (i.e. the interlocutor) and as such it is processed by the dialogue model that represents the patient. Before the dialogue model can begin processing the doctor's conversational behavior, the *values* of the `behavior's properties` need to be identified and described. The different types of properties are described in section [6.1] and the values are obtained from the two analyses described in chapter [5]. Based on the information described in chapter [5] and section [6.1], the following *values* are ascribed to the `properties of the conversational behavior` in this dialogue turn (i.e. step 1.1)):

a) `Intended effects:` *IntEff(salutation).*
IntEff(inform(D is present)).

b) `Expected effects:` *ExpEff(BEL$_P$(D is present), 1.0).*
ExpEff(ToPerform$_P$(BEH(expressive(greeting))).
ExpEff(ToPerform$_P$(Return greeting)).

c) `Configuration of IS:` *INT$_D$(ToPerform$_P$(BEH(expressive(greeting)))).*
INT$_D$(ToPerform$_P$(BEH(acknowledge(D is present)))).
INT$_D$(BEL$_P$(D is present), 1.0).
BEL$_D$(D and P can exchange information), 0.8.
BEL$_D$(P is present), 1.0.

d) `Relation to other beh.:` *BEH(expressive(greeting))*, $\rightarrow$*(BEH(expressive (greeting))).*
BEH(expressive(greeting)), $\rightarrow$*(BEH(question)).*
BEH(expressive(greeting)), $\rightarrow$*(BEH(statement)).*
BEH(expressive(greeting)), $\leftarrow$(none)

e) `Emotions:` *EMO$_D$(Sadness), 0.3.*
In addition, the conversational behavior contains expres-

sions of seriousness and reluctance in the doctor's face and voice.

f) `Social relation:` *$TRUST_{DP}$, 0.6.*
$LIKING_{DP}$, 0.6.
$RAPPORT_{DP}$, 0.6.

g) `Semantic content:` *The doctor greets the patient, John.*

These `behavior properties` and their associated *values* form the input of the dialogue model representing the patient and are used to process the conversational behavior performed by the doctor in dialogue turn D2.

The first step of processing the conversational behavior is to group it into one of the Categories of the Behaviors-base. The matching functions (i.e. categorization rules) used in the categorization processes indicate that most of the values from the first four properties (i.e. `the intended effect`, `the expected effect`, `the configuration of the internal state` and `the relation to other conversational behaviors`) match the values of the category properties that determine the category Greeting. Subsequently, the conversational behavior performed in dialogue turn D2 is placed in the Greeting category (i.e. step 1.2)).

During the second step in the process-cycle, the patient's internal state (which is represented in the state-bases of the dialogue model) is updated. The input for the various updating processes consists of several of the interpreted and described property values of the conversational behavior (i.e. for properties e), f) and g)) combined with the values of the category properties (i.e. for properties a), b), c), and d)).

As shown in chapter [6], the dialogue model consists of five bases in which the patient's internal state is represented (i.e. Dialogue state-base, Belief-base, Desire-base, Emotion state-base and Social relation-base). As a result, the second step in the process-cycle is divided into five sub-steps; one for each individual base.

Step 2.1) Updating the Dialogue state-base:

In the first sub-step, the dialogue model updates the elements in the Dialogue state-base. The information contained in category property d) of the Greeting category is used by the dialogue state updating processes to add the `behavior type` of the doctor's behavior to the sequenced list in the Dialogue state-base together with the appropriate attribute:

- *$Performed_D$(BEH(expressive(greeting)))*

In addition, the dialogue state updating processes also compare the *dialogue management rules* to the information contained in category property d) to determine which behavior types the response behavior can have. The response behavior can have the following behavior types, presented in degree of appropriateness: *expressive (greeting), question* and *statement*, as indicated by the $\rightarrow$ operator. In other words, the dialogue state updating processes determine that the most appropriate response should have the behavior type *expressive(greeting)*. This is in the Dialogue state-base as:

- *ToPerform$_P$(BEH(expressive(greeting)))*

Step 2.2) Updating the Belief-base:

As well as updating the dialogue state, the dialogue model also uses values of both the behavior properties and the category properties to update the elements in its Belief-base. This is the second sub-step in updating the patient's internal state. The belief formation and updating processes use the `the intended effects`, `the expected effects` and `assumptions` about the doctor's underlying internal state features of the conversational behavior to construct new **beliefs** or alter existing ones. In addition, the displays of the doctor's emotions are also taken into account during these processes. The belief formation and updating processes are realized by applying the *belief formation rules* on the values of the (behavior or category) properties and the additional information provided by the Social relation-base and the Dialogue state-base.

In response to the doctor's conversational behavior in dialogue turn D2, the dialogue model constructs the following new beliefs and, as they do not contradict with any already existing beliefs, stores them in the patient's Belief-base:

1. BEL$_P$(BEL(exp)$_D$(*ToPerform$_P$(BEH(expressive(greeting)))*), 0.8), 0.7.
2. BEL$_P$(BEL(exp)$_D$(ToPerform$_P$(Return greeting)), 0.8), 0.7.
3. BEL$_P$(BEL(exp)$_D$(BELP(D is present), 1.0), 1.0), 0.7.
4. BEL$_P$(INT$_D$(*ToPerform$_P$(BEH(expressive(greeting)))*)), 0.7.
5. BEL$_P$(INT$_D$(*ToPerform$_P$(BEH(acknowledge(D is present)))*)), 0.7.
6. BEL$_P$(INT$_D$(BEL$_P$(D is present), 1.0)), 0.7.
7. BEL$_P$(BEL$_D$(D and P can exchange information), 0.8), 0.7.
8. BEL$_P$(BEL$_D$(P is present), 1.0), 1.0.
9. BEL$_P$(EMO$_D$(Sadness), 0.3), 0.7.

The `expected effects` ascribed to the conversational behavior (described by the values of property b)) enable the dialogue model representing the patient to make assumptions about which expectations the doctor might have formed after performing his conversational behavior. The first three **beliefs** represent these assumptions about the doctor's expectations. The formation of the first two **beliefs** is supported by information provided by the Dialogue state-base. This information indicates that the most appropriate response behavior has the behavior type *expressive(greeting)* and should ideally be selected from the Return greeting category.

The dialogue model forms **beliefs** about the doctor's **intentions** and **beliefs** based on the assumption of the internal state configuration that caused the doctor to perform his conversational behavior (i.e. category property c)). The **beliefs** formed by the patient about the doctor's assumed underlying internal state are represented as **beliefs** four to eight. These do not include beliefs about the doctor's expectations. This is because the doctor's **expectations** (i.e. special beliefs) are formed after his performance of the conversational behavior, rather than being part of the preconditions.

The ninth **belief** (i.e. about the doctor's emotion) is formed by the dialogue model as a result of processing property e) of the conversational behavior. Because this

belief is based directly on the observation and the interpretations of the `displays of emotion` expressed in the conversational behavior, the belief formation and updating processes use the value of one of the behavior properties (i.e. property e)) instead of any category properties.

The certainty values assigned to the new **beliefs** are based on the fact that the patient trusts the doctor and the above average probabilities that the beliefs hold true. The patient's **degree of trust** is provided to the belief formation and updating processes by the Social relation-base. The values for the social elements are inserted manually into the dialogue model beforehand. Additional information about the social relations between the two interlocutors is obtained from any observed `displays of social aspects` (i.e. property f)) and the `semantic content` (property g)).

After performing this initial formation of **beliefs**, an iterative application of the belief formation and updating processes is executed by the dialogue model. As a result, several of the patient current (including his new) **beliefs** are updated, adopting some of the doctor's assumed **beliefs** as his own. The second application of the belief formation and updating processes produces the following new **beliefs**:

10. BEL_P(D and P can exchange information), 0.7.
11. BEL_P(P is present), 1.0.

The last newly formed belief is somewhat strange as P is obviously aware he himself is present at the conversation. This is an unforeseen flaw is the mechanism of the dialogue model, but in practice it shouldn't be problematic.

Step 2.3) Updating the Desire-base:

In addition to forming new **beliefs** and updating current ones, the dialogue model also makes adjustments to the patient **intentions**. The initial intention that the patient adopted was his **desire** to go home (derived from the introductory dialogue turn P1). After interpreting the doctor's conversational behavior, the patient shifts his **intention** to acknowledging the doctor's presence and to perform a conversational behavior from the Return greeting category. He does so because the new **beliefs** he has formed (based on the doctor's conversational behavior) make it more likely and more appropriate for him to achieve the **desires** to acknowledge the doctor's presence and to perform a greeting in return than the **desire** to go home. Note that the patient retains his **desire** to go home, but he isn't actively pursuing it at that time.

The patient's **desire** to acknowledge the presence of an interlocutor (i.e. DES_P(*ToPerform*$_P$*(BEH(acknowledge(D is present)))*)) was already present in his Desire-base at the start of the conversation, as it is a primary premise which underlies all behavior from the Greeting and Return greeting categories. The **desire** to perform a return greeting stems from a sense of politeness and how to properly hold a conversation.

By applying the *intention updating rules* to the information provided by the dialogue model's Belief-base about the newly formed **beliefs**, the intention updating processes make the following changes to the elements in the Desire-base:

The **desire** DES_P(*ToPerform*$_P$*(BEH(acknowledge(D is present)))*) is updated into the **intention** INT_P(*ToPerform*$_P$*(BEH(acknowledge(D is present)))*) while the

desire DES_P(ToPerform$_P$(Return greeting)) is updated into the **intention** INT_P(ToPerform$_P$(Return greeting)). At the same time the **intention** INT_P(go home) is updated to the **desire** DES_P(go home).

Step 2.4) Updating the Emotion state-base:

Sub-step four in the internal state updating step of the dialogue model's process-cycle consist of updating the elements contained in the Emotion state-base. Updating these elements is the result of the dialogue model performing an emotion appraisal of the conversational behavior in dialogue turn D2. In order to emotionally appraise the doctor's behavior, the dialogue model needs to determine what *values* for the ***appraisal variables*** can be assigned. The appraisal variable values are determined by evaluating the `semantic content` of the doctor's conversational behavior (i.e. behavior property g)) with respect to different elements in the internal state of the patient. These elements include the patient's newly formed **beliefs**, his **desires** and **intentions** and any **expectations** (i.e. special **beliefs**) the patient had formed previously, for example about the outcome of his earlier conversational behaviors. As a result of this evaluation, the following *appraisal variable values* are established with respect to the doctor's conversational behavior:

- ***Desirability-for-self:*** No relevant values are found
- ***Desirability-for-other:*** No relevant values are found
- ***Status:*** *Unconfirmed*
- ***Likelihood:*** No relevant values are found
- ***Praiseworthiness:*** No relevant values are found
- ***Appeal:*** *Liking*

Because the evaluation of the conversational behavior did not yield many *appraisal variable values*, it is difficult to properly determine how the conversational behavior affects the emotions of the patient. There is not enough information to elicit an emotional response to the doctor's conversational behavior. As a consequence, the appraisal processes do no make any alterations to the emotion state of the patient at this time.

Step 2.5) Updating the Social relation-base:

Finally, the dialogue model checks if any of the elements in the patient's Social relation-base are affected by the `displays of the social aspects` (behavior property f)) and the `semantic content` (behavior property g)) of the doctor's conversational behavior. The interpreted `social displays` and the `semantic content` allow the dialogue model to make assumptions about how the doctor perceives the social relations between the himself and the patient. This information is used by the both the belief formation and updating processes (see sub-step two) and the social relation updating processes to alter the appropriate variables in the bases associated with them. Because the conversational behavior does not contain any `displays of social aspects` and the `semantic content` is rather limited in what it expresses, the social relation updating processes do not alter the values of any of the social states.

In addition to the `social displays` and the `semantic content`, the social relation updating processes also use other information to check if the patient's interpre-

tation of the social relations needs to be altered. This information consists of the patient's newly formed and updated **beliefs**, possible contradictions and any elicited **emotions**. As the new **beliefs** formed by the patient based on the doctor's conversational behavior do not contradict with **beliefs** already present in the patient's Belief-base, the dialogue model does not adjust the value associated with the social variable **trust**. Furthermore, because the conversational behavior has not elicited any emotions, the values of the social variables **liking** and **rapport** also remain unchanged.

After the cognitive processes in the dialogue model have constructed new elements and / or altered existing ones to update the patient's internal state, the dialogue model selects an appropriate conversational behavior for the patient to perform. Selecting and performing an appropriate response behavior is the third step in the process-cycle performed by the dialogue model.

When an appropriate response behavior has been selected by the behavior selection processes in the dialogue model the patient takes the turn in the conversation. In other words, the patient becomes the speaker and the doctor becomes the listener. As a result, the dialogue shifts into the next dialogue turn, P2.

Step 3: Selecting and performing the conversational behavior in dialogue turn P2:
Hello doctor Rupert.

The third step in the process-cycle consists of two sub-steps, namely the selection of an appropriate response behavior (step 3.1)) and subsequently the performance of that behavior (step 3.2)).

In the first sub-step, the behavior selection processes check if and how the elements from the five newly updated internal state bases (i.e. the Belief-base, the Desire-base, the Emotion state-base, the Social relation-base and the Dialogue state-base) are related to the values of the four properties that define the categories in the Behaviors-base. The higher the number of elements that are related to property values of a particular category, the more likely it is that the dialogue model will select a conversational behavior from that category to perform as its response behavior. In our example, the following values of the four category properties correspond to elements in the patient's internal state bases.

Category property a)

The majority of the conversational behaviors in the Return greeting category have *IntEff(salutation)* as value for the `intended effect` behavior property. As a result, the value of category property a) is also IntEff(salutation). However, this `intended effect` is also the value of category property a) of the Greeting category.

The patient's **intentions** that hold at this point, i.e. INT_P(ToPerform$_P$(Return greeting)) and INT_P(*ToPerform$_P$(BEH(acknowledge(D is present)))*), are two possible **intentions** that can be associated with the `intended effect` *IntEff(salutation)*. Based on this information alone, the dialogue model could select an appropriate response behavior from either one of these two categories. However, the first **intention** indicates that the patient intends to perform a conversational behavior from the Return greeting category. As a result, conversational behaviors from that category are

preferred over conversational behaviors from the Greeting category.

Category property b)

As the patient's Belief-base does not yet contain any **expectations**, no relation can be established between the `expected effect` property (i.e. property b)) of any of the categories and the elements that make up the internal state of the patient. The patient will form new **expectations** only after performing the selected conversational behavior.

Category property c)

By comparing the newly updated elements with the configurations that make up the values of property c) for each of the categories, it was found that the most similarities existed between the configuration of underlying internal state features of the Return greeting category and the `configuration of the internal state elements` currently present in the dialogue model. The following internal state features are assumed to underlie the behaviors in the Return greeting category and thus form the configuration value of property c) of that category:

- P has the **belief** that D just performed a conversational behavior from the Greeting category.
- P has the **intention** to perform a conversational behavior from the Return greeting category.
- P has the **intention** to perform a conversational behavior with the behavior type *`expressive(greeting)`*.
- P has the **belief** that D is present.
- P has the **intention** to increase the value of D's perception of the **degree of rapport** between the two interlocutors. In other words, make D feel the **rapport** between him and P has been increased

This configuration corresponds to the following elements in the dialogue model's internal state bases, respectively in the Desire-base (the **intention**) and in the Belief-base (the two **beliefs**):

- BEL_P(Performed$_D$(Greeting)), 1.0
- INT_P(ToPerform$_P$(Return greeting))
- BEL_P(D is present), 1.0

The remaining two values of the configuration property, i.e. the **intention** to perform a *`expressive(greeting)`* `type of behavior` and the **intention** to increase the doctor's feeling of **rapport**, do not have corresponding elements in the patient's current internal state. If these elements had been present in the patient's Desire-base, the selection of a conversational behavior from the Return greeting category would have been even more founded.

Category property d)

During sub-step 2.1, the dialogue state updating processes had established and recorded that the conversational behavior performed by the doctor (i.e. the behavior

performed in D2) has the behavior type *expressive(greeting)*. This is indicated in the Dialogue state-base by the Performed_D operator.

In addition, the dialogue state updating processes have also determined that the behavior type of the response behavior that is most appropriate should be *expressive(greeting)*, which is represented in the Dialogue state-base by the ToPerform_P operator.

When the behavior selection processes take into account the behavior type recommended by the Dialogue state-base, the categories from which an appropriate response behavior can be selected is reduced to the Greeting and Return greeting categories. This is determined by comparing the value of the ToPerform_P operator (in the Dialogue state-base) with the values of category property d).

After comparing the current elements in the patient's internal state bases with the category property values each available category, the behavior selection processes select a behavior from the category that has the most property values and internal state elements in common. The specific behavior that is selected from the category is determined by the semantic content of the patient's **beliefs**. As a response to the doctor's conversational behavior from dialogue turn D2, the dialogue model selects the sentence *"Hello doctor Rupert."* from the Return greeting category to be performed.

Since the intensity values of the elements in the patient's emotion state and social state are not sufficiently high, they do not affect the form of the conversational behavior nor are any displays or expressions included during the performance.

When the dialogue model performs a conversational response behavior with the *expressive(greeting)* behavior type, the following functions apply with respect to the relations to behavior types of related conversational behaviors:

- *BEH(expressive(greeting))* in P2, $\rightarrow$*(BEH(statement))* in D3
- *BEH(expressive(greeting))* in P2, $\rightarrow$*(BEH(question))* in D3
- *BEH(expressive(greeting))* in P2, $\leftarrow$*(BEH(expressive(greeting)))* in D2

Related conversational behaviors entail those behaviors that precede or follow the patient's behavior in P2. These functions are the value of property d) of the patient's behavior and are used by the belief formation and updating processes to form **expectations** about the possible behavior type of the behavior the doctor might give in response. The following **expectations** are formed and stored in the patient's Belief-base:

- $\text{BEL(exp)}_P(\text{Increase}_D(\text{RAPPORT}_{DP}))$, 0.7
- $\text{BEL(exp)}_P($*(ToPerform$_D$(BEH(question)))* or *(ToPerform$_D$(BEH(statement)))*$)$, 0.7

The first **expectation** is that the patient expects his polite return greeting to strengthen the doctor's feeling of **rapport** between them. The second **expectation** is that the doctor is going to ask him a question or is going to tell him something. In addition, as a result of performing the selected conversational behavior, the dialogue state updating processes will change the ToPerform_P attribute in the dialogue state base to a Performed_P attribute and add it to the sequenced list.

After the patient has selected and performed an appropriate response behavior, the process-cycle has returned to its starting position, where it will wait for a new behavior that it can process. Once a new conversational behavior is perceived and interpreted, the process-cycle will again be applied by the dialogue model.

As a response to the conversational behavior performed by the patient in dialogue turn P2, the doctor performs a new conversational behavior, thereby switching the dialogue turn again and continuing the dialogue. The following sentence is uttered by the doctor in dialogue turn D3:

Step 1) Processing the conversational behavior in dialogue turn D3:
I'm sorry to hear the news. We were hoping that surgery and a course of chemo would cure this tumor.

Based on the analyses presented in chapter [5] and the representations explained in chapter [6], the various aspects of the conversational behavior performed in dialogue turn D3 can be interpreted (step 1.1)). As a result, the following *values* are assigned to the properties of the conversational behavior:

a) Intended effect: *IntEff(inform(EMO_D(Sympathy), 0.4.)).*
IntEff(inform($BEL(exp)_D$(surgery and chemo would cure the tumor))).

b) Expected effect: *ExpEff(Decrease$_P$(EMO_P(Sadness))).*
ExpEff(Increase$_P$($RAPPORT_{PD}$)).

c) Configuration of IS: *INT_D(BEL_P(EMO_D(Sympathy), 0.4), 0.8).*
INT_D(BEL_P($BEL(exp)_D$(surgery and chemo would cure this tumor), 0.7), 0.8).
INT_D(Decrease$_P$(EMO_P(Blame(doctor)))).
INT_D(Decrease$_P$(EMO_P(Sadness))).
INT_D(Increase$_P$($RAPPORT_{PD}$)).
BEL_D(tumor is not cured), 1.0.
BEL_D(EMO_P(Sadness), 0.8), 0.8.

d) Relation to other beh.: *BEH(statement), $\rightarrow$ (BEH(statement)).*
BEH(statement), $\rightarrow$ (BEH(question)).
BEH(statement), $\leftarrow$ (BEH(statement)).
BEH(statement), $\leftarrow$ (BEH(suggestion)).
BEH(statement), $\leftarrow$ (BEH(answer)).
BEH(statement), $\leftarrow$ (BEH(expressive(greeting))).
BEH(statement), $\leftarrow$ (BEH(acknowledge)).
BEH(statement), $\leftarrow$ (none).

e) Emotions: *EMO_D(Sadness), 0.4.*
EMO_D(Sympathy), 0.4.

f) Social relation: *$TRUST_{DP}$, 0.6*
$LIKING_{DP}$, 0.6
$RAPPORT_{DP}$, 0.6

g) `Semantic content:` The doctor expressing sympathy.
The doctor expected that surgery and chemo would cure the tumor.

During the first step of the process-cycle, the dialogue model compares the *values* of the first four behavior properties with the values of the corresponding category properties. Based on the similarities between the values, the conversational behavior is grouped into both the Inform about bad situation category and the Express sympathy category (step 1.2)). Grouping into more than one category is the result of the `behavior property` *values* corresponding to a majority of the category property values of each determined category.

After the doctor's behavior has been categorized, the process-cycle in the dialogue model will update the patient's internal state. The updating processes result in the following new and altered elements:

Step 2.1) Updating the Dialogue state-base:

By using the values from property d) of both categories in which the doctor's conversational behavior is grouped, the dialogue state updating processes determine that the behavior has the `behavior type` *BEH(statement)*. As a result, the dialogue state updating processes assigns the Performed_D attribute to the behavior type of the doctor's conversational behavior and adds it to the sequenced list in the Dialogue state-base:

- $Performed_D(BEH(statement))$

The dialogue state updating processes also determines that the doctor's behavior can be answered by a conversational behavior with one of the following types: either another *BEH(statement)* or a *BEH(question)*. This suggestion is derived from the information about `behavior type` relations also contained in category property d) of both categories. Both behavior types receive the ToPerform_P attribute as they are place in the Dialogue state-base:

- $ToPerform_P(BEH(statement))$
- $ToPerform_P(BEH(question))$

The behavior selection processes use these suggested `behavior types` to make a pre-selection of the categories from which an appropriate response should be selected.

Step 2.2) Updating the Belief-base:

After the elements in the Dialogue state-base have been updated, the belief formation and updating processes update the Belief-base by forming new **beliefs** about the situation based on the doctor's conversational behavior in dialogue turn D3. In this case, the belief formation and updating processes use the category property values of two categories to construct the new **beliefs**, as the doctor's conversational behavior is grouped in both the Inform about bad situation category and the Express sympathy category. As a result, the following new **beliefs** are formed by the dialogue model:

1. $BEL_P(BEL(exp)_D(Decrease_P(EMO_P(Sadness))), 0.5), 0.8$.
2. $BEL_P(BEL(exp)_D(Increase_P(RAPPORT_{PD})), 0.5), 0.7$.
3. $BEL_P(INT_D(BEL_P(EMO_D(Sympathy), 0.4), 0.8)), 0.7$.
4. $BEL_P(INT_D(BEL_P(BEL(exp)_D$(surgery and chemo would cure this tumor), 0.7), 0.8)), 0.7.
5. $BEL_P(INT_D(Decrease_P(EMO_P(Blame(doctor)))), 0.7$.
6. $BEL_P(INT_D(Decrease_P(EMO_P(Sadness)))), 0.8$.
7. $BEL_P(INT_D(Increase_P(RAPPORT_{PD}))), 0.7$.
8. $BEL_P(BEL_D$(tumor is not cured), 1.0), 0.8.
9. $BEL_P(BEL_D(EMO_P(Sadness), 0.8), 0.8), 0.7$.
10. $BEL_P(EMO_D(Sadness), 0.4), 0.7$.
11. $BEL_P(EMO_D(Sympathy), 0.4), 0.7$.

The first two of the newly formed **beliefs** describe the patient's thoughts about what the doctor anticipates to be the consequences of his conversational behavior. According to the patient, the doctor expects that his behavior will ease the patient's **Sadness** and builds up the degree of **rapport** between himself and the patient.

The next five elements represent what the patient thinks the doctor wants to achieve with his behavior. The patient **believes** the doctor **intends** to let the patient **believe** he (i.e. the doctor) is sympathetic to the patient's situation and that he expected that the surgery and chemo therapy would cure the tumor. In addition, the patient **believes** the doctor wants to shift the **Blame** for the patient's situation away from himself, comfort the patient (i.e. decrease the patient's feelings of **Sadness**) and increase the **rapport** between them.

The processing of the conversational behavior also causes the patient to form the **beliefs** that the doctor **believes/knows** that the tumor is not cured and that the patient is feeling sad, as is indicated by new **beliefs** 8 and 9.

By using the value of behavior property e) the belief formation and updating processes form beliefs 10 and 11. These last two **beliefs** indicate that the patient believes the doctor is experiencing respectively **Sadness** and **Sympathy**.

After the dialogue model has formed these twelve new **beliefs**, a second iteration of the belief formation and updating processes is executed to allow the patient to adopt some of the doctor's **beliefs** as his own. As a result the following **belief** is formed by the *belief updating rules*:

12. BEL_P(tumor is not cured), 0.8

However, previously the patient has formed the **belief** ***BELD(tumor is cured), 0.8***, based on what the surgeon has told him. As these two **beliefs** contradict each other and their certainty values are the same, both **beliefs** are included in the Belief-base but their certainty values are halved to 0.4. This represents the patient's uncertainty whether or not the tumor is cured.

Furthermore, the conversational behavior performed by the doctor has fulfilled one of the two **expectations** the patient had formed after selecting and performing his previous conversational behavior, namely the **belief** that the doctor was going to ask or utter a statement. The doctor's behavior didn't express that his view of the

rapport existing between himself and the patient had increase in intensity as a result of the patient's previous conversational behavior. As a result, that **expectation** of the patient is not yet confirmed and fulfilled.

Step 2.3) Updating the Desire-base:

While the dialogue model is updating the patient's Belief-base with new and altered **beliefs**, another step in the process-cycle affects the patient's **desires** and **intentions**. The **intentions** the patient previously had were fulfilled when he performed the conversational behavior in dialogue turn P2. As a consequence, those **intentions** were transformed back into **desires**. Now, new **intentions** need to be selected from the patient's **desires** in order to give rise to his next conversational behavior.

Because one of the newly formed **beliefs** contradicted with a **belief** already present in the Belief-base, the intention updating processes aims to select **desires** that focus on gathering information about the patient situation to adopt as the patient's new **intentions**. This aim is strengthened by information from the Dialogue state-base, which states that it is most appropriate to select a response behavior with either the *question* or the [*statement* behavior type. This is indicated by the $\text{ToPerform}_{\text{P}}$ attribute. **Desires** that aim to gather information were already included in the Desire-base of the dialogue model before the start of the conversation. The *intention updating rules* used by the intention updating processes update these **desires** into the following **intentions**:

- $\text{INT}_{\text{P}}(\text{ToPerform}_{\text{P}}(\text{Gather information(patient's health is good)})$
- $\text{INT}_{\text{P}}(\text{BEL}_{\text{P}}(\text{patient's health is good}), 1.0)$.
- $\text{INT}_{\text{P}}(\text{ToPerform}_{\text{D}}(\text{Inform about bad situation(patient's health is good)})$

These **intentions** indicate that the patient wants to select and perform a conversational behavior from the category Gather information in order to get to know (i.e. can form a **belief**) if the patient's health is good. At the same time, the patient **intends** to get the doctor to perform a response behavior that falls into the category Inform about bad situation.

Step 2.4) Updating the Emotion state-base:

At the same time the Belief-base and the Desire-base are being updated, the appraisal processes in the dialogue model evaluates how the semantic content of the doctor's conversational behavior (i.e. behavior property g)) and the doctor's underlying internal state features (i.e. category property c)) affect the patient's internal state features. As a result, the following *appraisal variable values* were determined:

- ***Desirability-for-self:*** *Undesirable.* The patient desires the situation in which the tumor is cured.
- ***Desirability-for-other:*** No relevant values are found
- ***Status:*** *Disconfirmed.* The patient thought ("expected") that the tumor was cured. The situation expressed in the conversational behavior disconfirms this expectation.

- ***Likelihood:*** *0.5.* Because the patient is not sure about whether the tumor is cured or not, his interpretation of the likelihood that it is cured is as big as the likelihood that it isn't cured.
- ***Praiseworthiness:*** No relevant values are found: The patient is not certain enough about the expressed situation to praise or blame the doctor.
- ***Appeal:*** *Disliking.* Based on the situation expressed

By comparing these values to the values predetermined for the different emotion types in the Emotion state-base, the appraisal processes determine that the emotions **EMO_P(Confusion)** and **EMO_P(Surprise)** are affected by the doctor's conversational behavior and that the intensity values of these elements need to be increased. This is subsequently done.

Step 2.5) Updating the Social relation-base:

Next, the social relation updating processes in the dialogue model evaluates how the features that define the social relation between the doctor and the patient are affected by the doctor's conversational behavior and by the situation that is expressed by it. The state of the situation is described by the behavior's `semantic content`, i.e. behavior property g).

Since the belief formation and updating processes produced a pair of contradicting **beliefs** and the emotional appraisal of the doctor's behavior resulted in the patient experiencing **Confusion** and **Surprise**, it is likely the patient is less inclined to accept what the doctor is saying than before.

As a result of this reduced inclination of acceptance, the intensity value of the **$TRUST_{PD}$** element in the patient's Social relation-base decreases. The intensity value of the **liking** element also decreases, even though the patient does not (yet) experiences **Blame** towards the doctor for giving him bad news. This decrease in **liking** caused by the fact that the situation expressed in the conversational behavior negatively affects the patient's wellbeing. Because the emotions **Sadness** and **Sympathy** are ascribed by the patient to the doctor, the patient's interpretation of the **rapport** between them is not negatively affected by the doctor's conversational behavior. After all, according to the patient the doctor tries to console him and feels **Sorry for him**.

After all internal state bases have been updated, the process-cycle in the dialogue model executes the third step, namely selecting and performing the patient's response behavior. As mentioned before, this elicits a shift in the turn-taking of the conversation. As a result, the selection and performance is considered part of dialogue turn P3.

Step 3: Selecting and performing the conversational behavior in dialogue turn P3:
What do you mean?

As part of step three of the process-cycle, the dialogue model checks if the elements in the updated internal state bases are related to the values of the categories' properties. Based on this comparison, the dialogue model is able to determine which category contains the most appropriate response behaviors. The following relations exist between the elements of the current (and newly updated) internal state and the values of each of the four category properties.

Category property a)

In order to determine what the `intended effect` property of the response behavior could be, the dialogue checks the Desire-base for the patient's **intentions**. One of the **intentions** that the patient's currently hold, (i.e. INT_P(ToPerform$_P$(Gather information(patient's health is good)))) can be related to the following `intended effect` of a conversational behavior:

- `IntEff(`*`question(patient's health is good))`*

The patient's other current **intention** is to get the doctor to provide him (i.e. the patient) with information about his health:

- INT_P(ToPerform$_D$(Inform about bad situation(patient's health is good)))

This **intention** can be related to the same `intended effect` of a possible response behavior as the first **intention**. Conversational behaviors that typically have *`question`* as a Dialogue act type and that are aimed at inquiring after information are most often contained in the Gather information category.

Category property b)

The **expectations** that are currently present in the patient's Belief-base do not relate to `expected effects` of any of the conversational behaviors he might perform, but rather related to the doctor's conversational behaviors and internal state. As a result of this, the patient's **expectations** are not used in the process of selecting an appropriate response behavior.

Category property c)

As a result of the comparison between the elements contained in the patient's internal state and the configurations of underlying internal state features that make up the values of property c) of the categories, the dialogue model has decided that a behavior from the Gather information category seems most appropriate as a response behavior. This is determined based on the fact that the value of property c) of this category matches best with the elements in the internal state. The generalized configuration that is assumed to underlie the conversational behaviors contained in the Gather information category consist of the following internal state statements:

- P has the **intention** to get D to perform a conversational behavior from the Inform about good/bad situation (*x*) category.
- P has the **intention** to know information x
- P has the **belief** that D knows information x
- P has the **expectation** that D will perform a conversational behavior from the Inform about good/bad situation (*x*) category

The two **intentions** in the configuration are similar to the current **intentions** contained in the Desire-base, i.e.

- INT_P(ToPerform$_P$(Gather information(patient's health is good)))

- INT_P(ToPerform$_D$(Inform about bad situation(patient's health is good)))

but only if the *x* variable is replaced with the value (patient's health is good).

If the *x* variable in the first **belief** is replaced with a value that indicates one of the two possible states of the patient's health, i.e. (patient's health is good) or (patient's health is not good), the behavior selection processes can relate this configuration feature to one of the **beliefs** currently present in the patient's Belief-base, namely:

- BEL_P(BEL_D(tumor is not cured), 1.0), 0.8.

Only the last feature of the configuration cannot be related to any of the elements in patient's current internal state. As mentioned, the patient's current **expectations** focus on whether the doctor's perception of the **rapport** between him and the patient increases or not. The underlying **expectation** of the configuration can thus not be matched to any **expectation** currently present in the Belief-base.

Category property d)

The elements in the Dialogue state-base inform the behavior selection processes that the most appropriate response behavior to select has either the `behavior type` *`BEH(question)`* or the `behavior type` *`BEH(statement)`*. This information, combined with the information obtained from the comparison of the value of category property a) with elements of the internal state (i.e. the `intended effect` that can be related to the patient's current **intentions**) enables the behavior selection processes to determine that the suggested `behavior type` *`BEH(question)`* is the most appropriate choice. The behavior type relations of a conversational behavior of the `behavior type` *`BEH(question)`* are:

- BEH(question), $\rightarrow$(BEH(answer)),
- BEH(question), $\leftarrow$(BEH(statement))
- BEH(question), $\leftarrow$(BEH(question))

These relations correspond with the functions that make up property d) of the Gather information category, strengthening the selecting of a conversational behavior from that category, in this case the sentence "what do you mean."

Possible alternative conversational behaviors that could have been selected if they were included in the Behaviors-base would be "What does this mean for my health?" or "Does this mean I'm still sick?" as both these questions inquire after the status of the patient's health.

After selecting the conversational behavior from the Behavior-base, the dialogue model augments its performance by including `displays` of the patient's **emotions** and/or **social dispositions**. As a result of the doctor's previous conversational behavior, the intensities of the patient's emotions of **Confusion** and **Surprise** are sufficiently high enough to cause them to be displayed. The emotions are expressed via the patient's `facial expression` and the `prosody of the utterance` performed in the selected conversational behavior.

The increased intensity of the social aspect of **rapport** manifest itself through an increased degree of politeness in which the behavior is performed.

Following the performance of the selected and adjusted conversational behavior, the belief formation and updating processes in the dialogue model form the following **expectations**:

- $BEL(exp)_P(BEL_D(BEL_P(\text{patient's health is not good}), 0.4), 0.9), 0.7$.
- $BEL(exp)_P(\mathit{ToPerform}_D(\mathit{BEH(answer)}), 0.7$.

The patient's first **expectation** expresses that as a result of the patient's question, the doctor will form the **belief** that the patient is unsure about his health (i.e. has a **belief** with a low certainty value, namely 0.4). This **belief** is aimed to let the doctor understand the `intended effect` of the patient's behavior, namely to inquire if the patient's health is good.

The second **expectation** is that the doctor will perform a conversational behavior that has the `behavior type` *`answer`*. This **expectation** stems from the dialogue management rules that represent adjacency pairs.

In addition to forming new **expectations**, the dialogue state updating processes will change the ToPerform_P attribute in the Dialogue state-base to a Performed_P attribute (for the $\mathit{BEH(question)}$) and adds it to the sequenced list. The dialogue state updating processes will delete the $\mathit{ToPerform}_P(\mathit{BEH(statement)})$ as this option has not been used.

7.2 Conclusions

In this chapter we presented a step-by-step example of the workings of the dialogue model. It illustrated how the conversational behaviors of several consecutive dialogue turns were either processed or selected and performed.

One remark that can be made in response to this example, is that is seems like the forming and updating of the internal state elements as a result of the processing of an observed conversational behaviors is done in such a way that it unmistakably leads the selection of the consecutive conversational behavior in the dialogue.

However, considerable effort was put into constructing the processes in such a way that such direct relations were avoided as much as possible. It is possible that the precise premises for selecting a conversational behavior from a particular category need to be tweaked and calibrated, but such adjustments would presumably be small.

We argue that the mentioned interpretation of the workings of the dialogue model can be explained by the limited amount of data that is contained in the dialogue model's internal state bases in the course of this example. If the dialogue model would be implemented in a functioning embodied conversational agent, then after having held multiple conversations, the dialogue model would contain more internal state features which facilitate the selection of the best possible response behaviors.

Note that this relation is directly proportional, i.e. the more conversations the agent has had, the better and more precise the selection processes in the dialogue model can be executed.

Another conclusion we draw from the example described in this chapter is that the various processes in the dialogue model are more extensive and complex than is apparent at first glance. The fact that we needed sixteen pages to describe the processing and selection of five simple dialogue turns illustrates this adequately.

8

Chapter 8: Conclusions

Our primary motivation for performing the study presented in this thesis was to find out how the processing and selection of the behaviors performed in natural conversations functions and how knowledge of these cognitive processes could be used in the construction of complex dialogue systems. In order to find answers to these questions, we analyzed a large variety of linguistic and psychological theories, methods and models focusing on various aspects of conversational behaviors and the mechanisms involved in processing them. Moreover, we performed a study on the properties of conversational behaviors. This resulted in the construction of a cognitive dialogue model, which is based on the conclusions drawn from this analysis and the performed study. This dialogue model can function as an architecture for a dialogue system.

In this final chapter, we present the overall conclusions we have drawn from our analysis of the theories and methods, our own studies into the workings of conversational behavior and the dialogue model we subsequently designed. Section [8.1] provides answers to each of the research questions we posed in chapter [1]. In section [8.2] we discuss the contributions and the limitations of this thesis. In section [8.3], we give some suggestions on how to proceed in the future based on the study presented in this thesis.

8.1 Research conclusions

Research question 1:
1) What determines the meaning and purpose of a conversational behavior?

In chapter [2], we showed that the main purpose of a conversational behavior is to bring about a particular situation or event. This includes, but is not limited to, a new state of the environment, a change in the thoughts or feelings of the interlocutor (i.e. the listener), a change in the thoughts or feelings of the speaker himself, or the performance of a particular action by either the listener or the speaker.

The meaning of a conversational behavior depends on the context in which it is performed and by what the speaker intends to achieve by performing it. In chapter

[4] we demonstrate that the situation or event that the speaker wants to achieve is determined and defined by the speaker's current internal state. Consequently, we argue that each conversational behavior can be linked to a particular configuration of the speaker's internal state.

Based on this notion, we make a distinction between the `properties of a conversational behavior` and the **features in the internal state** of the speaker. For example, the objective or goal of a conversational behavior can be described by the `intended effect` behavior property, whereas the objective of the speaker is described as his **intention** (i.e. an internal state feature). If we can attribute an intended effect to a conversational behavior, then we can also make the assumption that the listener can derive the associated part of the speaker's internal state (i.e. his intention) from that behavior.

Because it is much easier to derive the intended effect from a conversational behavior than to ascribe an intention to the speaker, intended effects have been the focus of many linguistic theories and models. In chapter [2] we described how the theories of Austin and Searle define this behavior property and show how different intended effects can be clustered. As many dialogue act theories build on the theories of Austin and Searle, these were explained in section [2.4]. One theory we focus on in particular, is the DIT++ taxonomy of dialogue acts.

The DIT++ taxonomy provides us with an extensive overview of definitions describing the intended effects (called the **communicative function** in the DIT++ taxonomy) of behaviors one might observe in a conversation. However, the intended effects are defined (with respect to the internal state) rather vaguely in the DIT++ taxonomy, using expressions such as what a speaker wants, assumes or knows.

Because we are more interested in the internal state features that are associated with a conversational behavior than with a description of its properties, using the DIT++ taxonomy does not seem to be sufficient. We conclude that the intended effects defined in the DIT++ taxonomy need to be extended so they describe more clearly how the properties of a conversational behavior are related to features of the speaker's internal state. Not only does this allow us to make assumptions about the internal state of the speaker, but also about how a conversational behavior may influence the internal state of the listener. We do so through our own taxonomy of conversational behaviors described in section [6.1].

Research question 2:
2) Which aspects of a conversation (within a particular domain) need to be taken into account when determining what an appropriate response behavior is?

The second research question is answered in chapter [3] for the domain of medical bad news conversations. In study of various models and techniques focusing on holding bad news conversations, we found that what is appropriate as a response behavior depends for a large part on the situation the interlocutors are in and their dispositions with regard to that situation. The content of the internal state plays an important role, not only for processing the observed conversational behavior and evaluating the situation, but also for selecting the most appropriate response behavior.

Within the domain of bad news conversations in a medical environment, the selection of the patient's response behaviors is for a large part guided by two cognitive traits. The first trait is the coping strategy of the patient. This coping strategy acts as a filter for all response behaviors that do not correspond to the patient's disposition and approach to the bad news situation. The second trait consists of the emotions the patient is experiencing. When one or more emotions in the emotion state have attained a high enough intensity, they override the normal selection of a response behavior and instead will cause the patient to select and perform a response behavior that expresses this emotion.

For the doctor, the aim to balance the patient's state of emotions appears to significantly influence the processes of his behavior selection. The same holds for his aim to bring across as clearly as possible what he is saying and more importantly what that means for the patient.

Research question 3:
3) Which cognitive features and processes are involved in processing conversational behaviors and selecting an appropriate response behavior?

In order to determine which internal state features and processes are used during the processing and selection of conversational behaviors performed in a natural dialogue, we studied various psychological theories and models that focus on such topics. Based on this study, presented in chapter [4], we concluded that in order to provide a realistic description of the processing involved in a natural dialogue at least the following types of internal state features should be used: **Beliefs** (to express the knowledge possessed by the speaker), **Goals / Desires** (to indicate all the situations that the speaker would like to be the case), **Intentions** (to express which situation the speaker is actively trying to bring about), **Emotions** (to indicate how the current situation is influencing the speaker's feelings) and **Social features** (to indicate the social relation that exists between the speaker and the listener). Furthermore, the internal states of the persons conducting a conversation also contain various mechanisms (i.e. processes) that carry out the formation and manipulation of the features of the internal state.

Such internal state features can be ascribed to the persons holding the conversation by examining the conversational behaviors they perform. The properties of behaviors can be determined directly from a conversation, but one can only make assumptions about the internal state features that hold in the speakers internal states. To determine the relation between behavior properties and internal state features, two analyses of the conversational behaviors that were expressed in a bad news conversation were performed. The setup and the findings of these two analyses were presented in chapter [5].

From the analyses, we concluded that beliefs and intentions are derived more often from interpreting the dialogue act type, the behavior type and the semantic content properties of a conversational behavior than from interpreting any explicit displays. Furthermore, we found that emotions are displayed much more explicit than beliefs and intentions and thus can be perceived more clearly. As a result, ascribing

emotions to a speaker's internal state based on interpretation of the conversational behavior can be done more easily.

In order to show how the processing of conversational behaviors functions, we have constructed a cognitive dialogue model using the determined internal state features. We aimed to construct a dialogue model that is capable of taking into account the internal state of an interlocutor when selecting its own behaviors. This dialogue model is presented in detail in chapter [6]. The core of our cognitive dialogue model is based on the Belief-Desire-Intention (**BDI**) model of human practical reasoning (Bratman, 1987). According to Bratman's model the use of beliefs, desires and intentions is essential for the processing of situations that might be brought about in the future.

We decided to use this model as a basis also because it relies on folk psychology. Folk psychology is the natural capacity to predict and explain the behavior and mental state of other people, for example the interlocutor in a conversation. It uses every day linguistic to describe features and processes that are involved in the processing of behaviors (including conversational behaviors).

By examining the properties of his conversational behaviors to form assumptions about the interlocutor's internal state features, the dialogue model can determine which response behavior is most appropriate given the situation and which behaviors it needs to perform to alter features in the interlocutor's internal state. Furthermore, the BDI-model is also often used in agent-based cognitive models.

The studies presented in chapter [4] indicated that conversational behaviors in natural dialogue often contain numerous displays of emotions and/or social features. This especially holds for conversations in which bad news is being discussed (see chapter [3]). In order for our dialogue model to be fully able to process complex, natural conversations, we augmented the BDI model with emotion processing and social relation processing capabilities. We argue that the addition of these capabilities also allows our cognitive dialogue model to perform in a more human-like manner.

In sections [4.2] and [4.3] we discussed various models and theories that focus on the processes of emotion appraisal and coping. The study of these models and theories provided us with information on how emotions are related to conversational behaviors and to other internal state features. In particular, the emotion types in the OCC model describe clearly the premises that need to hold in a situation or event to lead to the experience of an emotion in the internal state.

The coping models presented in section [4.3] showed how coping strategies might influence the selection of conversational behaviors. We selected the most appropriate coping strategies found in the various models and combined them with the coping strategies described in chapter [3]. We believe that this selection is sufficient to represent the various dispositions a patient may have regarding the situation and the approaches he subsequently may follow to deal with them.

The theories we studied and presented in section [4.4] describe how the social relations that exist between interlocutors might influence the features and processes contained in their internal states. In addition, they also indicate how certain properties of conversational behaviors can be related to social state features of both the speaker and the listener. From this study we concluded that these features are in-

terconnected to almost all other features and processes of the interlocutors' internal states.

We only performed a small study of the various social aspects of conversational behavior and the social relations that exist between interlocutors. One of the conclusions we draw from chapter [4] is that additional research on social aspects needs to be performed in order for the contribution of the social relations features to the processes involved in the processing and selecting of conversational behaviors in our dialogue model to become more significant and more realistic.

Research question 4:
4) How can such cognitive features and processes be represented in a cognitive dialogue model?

As described in chapter [6], the features and processes that make up the internal state in the cognitive dialogue model are represented in a similar manner as the ones used in the BDI model. In the BDI model, the internal state features (i.e. beliefs, desires and intentions) are represented as different modalities contained in a formal modal logic. Because this form of representation provides a structured and transparent way of expressing the complex and often unclear thoughts and feelings involved in processing conversational behaviors, it was used to represent all internal state feature in the dialogue model. This uniformity also ensures continuity throughout the dialogue model's internal state and allows the processing rules to facilitate interactions between internal state features of different types.

The cognitive processes are represented as strict rules that form and manipulate internal state features. As we did not aim to construct a formal logic model, we omitted the logic operators and functions. Most of the processes in our dialogue model function by comparing and trying to match internal state features from different sources (i.e. feature bases). Often, a match between two features functions as a premise of condition for a rule to trigger.

For example, the dialogue model has a belief X in its internal state and a rule in its behavior selection rules-base that states that "if belief X holds; perform behavior Y". One of the behavior selection processes compares all the beliefs in the belief-base with the premise of the rule and because it finds a match the rule triggers and the associated behavior Y is performed.

Research question 5:
5) How does the process of selecting an appropriate conversational response behavior operate?

When selecting an appropriate conversational response behavior, a person normally takes into account the situation he wants to bring about (i.e. his **desires** and **intentions**), the knowledge he possesses (i.e. his **beliefs**) and how he is feeling with regards to the situation and his interlocutor (i.e. his **emotions** and **social disposition**). In addition, he takes into account what he thinks his interlocutor is thinking and feeling, and how his response behavior might influence his interlocutor's internal

state. (These latter thoughts are represented by assumptions and expectations the person may form).

As described in chapters [6] and [7], the cognitive dialogue model selects appropriate conversational response behaviors by examining the configuration of its current internal state features. It does so by using various *behavior selection rules*:

These rules check the desires and intentions in the dialogue model to see which situation the responding person (or in this case the dialogue system) wants to bring about and is aiming to achieve. The beliefs are checked to see which assumptions and expectations are stored in the dialogue model regarding the current situation and possible desired alternatives.

In addition, the dialogue system's disposition towards the current and desired situation is checked by examining the dialogue model's emotion state (i.e. which emotions is the dialogue system experiencing). Furthermore, the *behavior selection rules* check the system's perception of the social relations that exists between the interlocutors.

Next, the *selection rules* check the state of the dialogue which the dialogue model has kept track off. This ensures that the sequence of alternating conversational behaviors follows a proper course.

When the dialogue model's current configuration of internal state features is mapped out, the *behavior selection rules* try to match as many features from the configuration as possible to the selection premises of the conversational behaviors that are stored in the behavior-base. These selection premises state which features must hold in the internal state, before the conversational behavior can be selected and performed. The conversational behavior that has the most internal state features matching with its premises is selected subsequently by the rules.

However, as also explained in chapter [6] there are a few exceptions to this rule. One exception is that the general selection of an appropriate response behavior can be overridden when the intensity of one of the emotions is very high (i.e. it exceeds a certain threshold). When this is the case, the behavior selection process will automatically select a conversational behavior that expresses the associated emotion. If more than one emotion has a very high intensity, the selection process selects a behavior that expresses the emotion with the highest intensity.

Another exception to the general selection of a conversational response behavior is when the speaker is following a specific tactic while holding the conversation, e.g. in the situation of dealing with bad news this tactic is expressed in the person's coping strategy. When such a tactic or strategy is being followed, the behavior selection process functions as normal, but an additional premise is added for selecting a behavior, namely that the response behavior fits within the strategy.

8.2 Discussion

Although it is difficult to say much about the quality of our cognitive dialogue model without implementing it in a dialogue system, there are some remarks we would like to make regarding this topic.

The dialogue model has not been tested in a conversational situation, nor has it

been validated. One of the reasons for this is that the dialogue model on its own does not allow for any testing. For this a working dialogue system that uses our dialogue model as its design architecture is required. Such a dialogue system would at the least need to contain a speech recognition component and a text to speech component to be able to communicate verbally and vocally. In order to also be able to express and recognize displays of emotions and social aspects that use posturing, gestures and facial expression facial recognition, image processing and a complete functioning embodied conversational agent capable of forming facial expressions need to be included.

Because it was impossible to construct such an extensive system within the scope of this research, we were unable to test if the dialogue model would function as naturally and appropriately as a human being. However, since we constructed the components in the dialogue model based on specialized methods, techniques and models that have been tested and validated on their own, we expect our dialogue model to be quite solid as well.

Currently the dialogue model does not contain enough data to make actual choices when selecting response behaviors. However, since it is capable of storing perceived and interpreted conversational behaviors, deriving the properties and associated internal state features from the behavior, and (re-)using these conversational behaviors as possible response behaviors, this argument might quickly be countered when the model is implemented into a dialogue system.

In our opinion, one of the major contributions of our cognitive dialogue system is that it focuses on taking into account the internal state of the interlocutor when processing conversational behaviors and selecting the most appropriate response behaviors. We found that many existing dialogue models lack this ability and consequently come up short as models of natural behavior. In chapter [6], we have shown that it is possible to construct a behavior selection mechanism that not only incorporates the interlocutor's current state, but that also allows for responses to be selected that are understood to influence the interlocutor's internal state.

Furthermore, the ability to take into account the fact that emotions experienced by both the speaker and the listener may influence the listener's behavior is expected to help in creating more realistic dialogue systems. Attribution emotions such as **Blame** and **Praise**, and the emotion types **Sympathy** and **Empathy** play important roles in the construction of emotion focused dialogue systems. We found that these were missing in many of the studied emotion appraisal models.

The same holds for the inclusion of the speaker's disposition with respect to the social relation between him and the listener, although this has remained underdeveloped in our current dialogue model.

In addition to the contributions provided by the dialogue model, the categorization method for grouping conversational behaviors also provides suggestions on how internal state features can be ascribed to a speaker based on the extraction of the properties of a conversational behavior.

8.3 Future work

Although the cognitive dialogue model contains all the necessary components and processes to function, there are several points upon which the model can be extended and improved. In this section we make several recommendation on how this might be achieved in future work. We argue that there are two major points that need to be addressed to further the development of our dialogue model. The first one, is the representation of the social relation aspects. To produce a more human-like result from the dialogue model, this component needs to be extended.

Second, to be able to validate the cognitive dialogue model we need to implement it into a complete embodied conversational agent system. This way we can determine if the assumptions we made about how the processes function and if the representations of the internal state features are correct. Also if any flaws exist they will come to light and they can subsequently be fixed. In addition, by testing the dialogue model in practice, we can study how tweaking the variables in the model will affect the processing and performing of conversational behaviors. This will enable us to work out the functioning of the various coping strategies in more detail. Furthermore, by employing the dialogue model to hold actual conversations it can gather more data to be included in its internal state. This in turn should also improve the quality of the conversations the dialogue model holds, as it allows the dialogue model to select its behaviors from a broader range.

DIT++ Taxonomy

In this appendix, the definitions of the `Intended Effect behavior properties` (also known as the behavior's `communicative functions`) are presented, as they are obtained from the DIT++ taxonomy (Release version 5, May 2010) Found at:
http://dit.utv.nl[1].

THE DIT++ TAXONOMY:

General-Purpose Communicative Functions:
These are functions that can be applied to any kind of semantic content. They are often applied to information concerning the task or activity that motivates the dialogue, and in that case they form a dialogue act in the Task/Activity dimension. General-purpose functions can also be applied to content concerning the communication, in which case they form a 'dialogue control act'.

- *Information Transfer Functions:*
  The class of information-transfer functions consists of all those functions whose primary aim is to obtain or to provide information. The class falls apart into *information-seeking* and *information-providing* functions.
  - *Information-seeking functions:*
    All functions in this class have in common that the speaker wants to know something, which he assumes the addressee to know.
    So-called Check Questions carry the additional assumption that the speaker expects the answer to be that the proposition under discussion is true. Still more specifically, some check questions carry the additional assumption that the addressee's beliefs confirm the speaker's expectation (*Posi-Check*) or that they contradict this expectation (*Nega-Check*).
    - *Question:* The speaker S wants to obtain certain information, which he assumes the addressee A to possesses, and puts pressure on A to provide this information.
      - *Propositional Question* (a.k.a. *Yes/No-Question*)*:* The information that S wants to obtain is the truth value of a proposition *p*.
        - *Check Question:* S weakly believes that the proposition is true.

[1]Last accessed on 24-03-2014

- *Posi-check:* S weakly believes that A also believes that the proposition is true.
- *Nega-check:* S weakly believes that A believes that the proposition is false.

- *Set Question* (a.k.a. *WH-Question*)*:* The information that S wants to obtain is which elements of a certain set, described in the semantic content, have a certain property, as also described in the semantic content. S assumes that there is at least one such element, S assumes that A knows some or all of the elements of that set which have that property.
- *Choice Question* (a.k.a. *Alternatives Question*)*:* S wants to know which one from a list of alternative propositions, described in the semantic content, is true; S believes that one of the alternatives is true.

- *Information-providing functions:*
  All information-providing acts have in common that the speaker provides the addressee certain information which he believes the addressee not to know or not to be aware of, and which he assumes to be correct. The various subtypes of information-providing functions differ in the speaker's motivation for providing the information, and in different additional beliefs about the information that the addressee possesses.
  - *Inform:* Speaker S wants to make the information p that forms the semantic content of the inform act known to addressee A; S assumes that the information p is correct.
    - *Agreement:* S believes that A weakly believes the semantic content to be true.
    - *Disagreement:* S believes that A weakly believes the semantic content to be false.
      - *Correction:* S wants the semantic content, which he believes to be correct, to replace a belief by A that S believes to be incorrect.
    - *Answer:* S believes that A wants to possess the information which forms the semantic content of the Answer act.
      - *Confirm:* S believes that A weakly believes that the propositional content is true.
      - *Disconfirm:* S believes that A weakly believes that the propositional content is false.

- *Action Discussion Functions:*
  Action Discussion functions have a semantic content consisting of an action, and possibly also a description of a manner or frequency of performing the action. This frequency may be zero, so e.g. an Instruct to perform an action with frequency zero is the same as prohibiting that action, and committing oneself to perform an action with zero frequency is the same as committing oneself to not

perform the action.

- *Commissives:*
  S is committed to performing a certain action in a certain manner or with a certain frequency, possibly dependent on certain conditions, and possibly dependent on A's consent that S do so.
  - *Offer:* S is committed to perform the action in the manner or with the frequency, described in the semantic content, if A would like S to do so.
  - *Promise:* S is committed to perform the action in the manner or with the frequency, described in the semantic content; S believes the action to be beneficial for A.
    - *Threat:* S is committed to perform the action in the manner or with the frequency, described in the semantic content; S believes the action to be harmful for A.
    - *AddressRequest:* S assumes that A wants S to perform the action; S is committed to conditionally perform the action, with conditions (for instance concerning the manner or frequency of performing the action), described in the semantic content.
      - *AcceptRequest:* S is committed to unconditionally perform the action described in the semantic content.
      - *DeclineRequest:* S is committed to not perform the action described in the semantic content (i.e., S is committed to perform the action with frequency zero).
    - *AddressSuggestion:* S assumes that A believes that the action, described in the semantic content, is potentially promising for achieving a certain goal; S also assumes that A believes that S is able to perform the action (possibly together with A); S is committed to conditionally perform the action, with conditions (for instance concerning the manner or frequency of performing the action), described in the semantic content.
      - *AcceptSuggestion:* S is committed to unconditionally perform the action described in the semantic content.
      - *DeclineSuggestion:* S is committed to not perform the action described in the semantic content (i.e., S is committed to perform the action with frequency zero).
- *Directives:*
  S wants A to consider a certain action which A might carry out (possibly together with S), potentially wanting to put pressure on A to do so.
  - *Instruct:* S wants A to perform the action in the manner or with the frequency described in the semantic content; S assumes that A is able to do so.
    - *AddressOffer:* S believes that A is committed to perform the action de-

scribed in the semantic content dependent on S's consent that A do so.

- *AcceptOffer:* S wants A to perform the action described in the semantic content
- *DeclineOffer:* S wants A to not perform the action described in the semantic content (i.e., S wants A to perform the action with frequency zero).

- *Request:* S wants A to perform the requested action in the manner or with the frequency described, conditional on A's consent; S assumes that A is able to do so.
  - *Indirect Request:* S wants A to perform the requested action in the manner or with the frequency described, conditional on A's consent.
- *Suggestion:* S wants A to know that the action in the manner or with the frequency described in the semantic content, is potentially promising for achieving a certain goal, which either S believes A to have, or which is specified as part of the semantic content; S assumes that A (possibly together with S) is able to perform the action in the manner or with the frequency described.

Dimension-Specific Communicative Functions:
The DIT++ taxonomy specifies ten distinct dimensions in which the **Dimension-specific functions** are grouped. Nine of the ten dimensions contain dimension-specific communicative functions that handle the management of dialogues. These communicative functions are called **Dialogue control functions**. The tenth dimension, Task / Activity, contains dimension-specific communicative functions that express particular domain-related purposes of a dialogue act.

- *Domain-related Functions (Task/Activity Dimension):* are functions, expressible either by means of performative verbs denoting actions for performing activities in a specific domain, or by means of graphical actions such as highlighting, or pointing to something in a picture. For example:
  - Open Meeting, Suspend Meeting, Resume Meeting, Close Meeting (in meeting situations).
  - Bet, AcceptBet.
  - Congratulation, Condolence (in appropriate social situations).
  - Hire, Fire, Appoint,... (in a human resource management domain).
  - Show, Highlight, Point, List,... (for performing multimodal dialogue acts).

- *Dialogue Control Functions:*
  The functions of communicative acts that serve to create or maintain the conditions for successful interaction.
  - *Feedback Functions:*
    Feedback acts provide or elicit information about the processing of he previ-

ous utterance(s), where at least five levels of attending to an utterance and processing it are distinguished:

1. *attention*, i.e. paying attention to the dialogue partner sufficiently to fully enable the perception of the partner's contributions (e.g. listening, looking).
2. *perception*, i.e. the recognition of the auditive, visual, or tactile components of communicative behavior.
3. *interpretation*, i.e. the assignment of meaning to the recognized communicative behavior. In terms of dialogue acts, this is the assignment of semantic content and communicative functions to utterances.
4. *evaluation*, i.e. comparing the information that an utterance encodes, due to its communicative functions and semantic content, with what was already known. For instance, when a question was asked to which, according to the addressee, the questioner already knows the answer, then the addressee cannot accept the information conveyed by the question, as this would put him in an inconsistent belief state.
5. *execution*, also called 'application' or 'dispatch'. For instance, execution of a request or instruct is performing the requested or instructed action; execution of a question is gathering the information to answer; executing an answer is integrating its semantic content with the belief state.

- *Auto-feedback functions (Auto-feedback dimension)* are about the speaker's own attention and processing of an utterance in the addressee's last turn. Auto-Feedback functions are intended to signal that the processing of the utterance in question failed at a certain level or was successful up to a certain level, ranging from attending via perceiving, understanding, and evaluating to doing something with the result of the processing at all these levels ("execution"). (More articulate feedback acts, signaling or requesting help for a specific processing problem, are constructed with general-purpose functions and a specific processing-related semantic content.)
  - *AutoPositive (= Unspecified Positive Auto-Feedback):* S successfully processed the previous utterance, but provides no information about the level(s) of processing being reported.
  - *AutoNegative (= Unspecified Negative Auto-Feedback):* S was unsuccessful in processing the previous utterance, but provides no information about the level(s) of processing being reported.
  - *ExecPositiveAutoFeedback (= Overall Positive Auto-Feedback):* S's perception, interpretation, evaluation, and execution of the previous utterance were successful.
  - *EvalPositiveAutoFeedback:* S's perception, interpretation, and evaluation of the previous utterance were successful.
  - *InterprPositiveAutoFeedback:* S's perception and interpretation of the previous utterance were successful.

- *PerceptPositiveAutoFeedback:* S's perception of the previous utterance was successful.
- *AttentPositiveAutoFeedback:* S is paying full attention.
- *ExecNegativeAutoFeedback:* S's perception, interpretation, and evaluation of the previous utterance were successful, but that he encountered a problem in applying the information from that utterances (for example, S was unable to carry out an instruction, or to find the information needed for answering a question).
- *EvalNegativeAutoFeedback:* S encountered a problem in evaluating the semantic content of the previous utterance (for example, the utterance provided information that is in conflict with information already available to S).
- *InterprNegativeAutoFeedback:* S's perception of the previous utterance was successful, but he encountered a problem in trying to assign an interpretation to the utterance (for example, S was unable to make sense of the semantic content).
- *PerceptNegativeAutoFeedback:* S's perception of the previous utterance encountered a problem (S did not hear the utterance well, or was unable to read it).
- *AttentNegativeAutoFeedback (= Overall Negative Auto-Feedback):* S did not pay (full) attention to the previous utterance (e.g., S did not listen carefully).

• *Allo-Feedback Functions (Allo-Feedback Dimension)* are about the speaker's beliefs about the addressee's attention and processing of an utterance in the speaker's last turn. A distinction is made in Allo-Feedback Functions between functions that provide (give) feedback and functions that elicit feedback:

Feedback-Giving Functions: S wants to provide information concerning his beliefs about A's processing of S's previous utterance.

- *AlloPositive (= Unspecified Positive Allo-Feedback):* S believes that A successfully processed the previous utterance, but provides no information about the level(s) of processing being reported.
- *AlloNegative (= Unspecified Negative Allo-Feedback):* S believes that A was unsuccessful in processing the previous utterance, but provides no information about the level(s) of processing being reported.
- *Overall Positive Allo-Feedback:* S believes that A's execution of S's previous utterance was successful.
- *Execution Negative Allo-Feedback:* S believes that A's execution of S's previous utterance was unsuccessful.

- *Evaluation Positive Allo-Feedback:* S believes that A's evaluation of S's previous utterance was successful.
- *Evaluation Negative Allo-Feedback:* S believes that A's evaluation of S's previous utterance was unsuccessful.
- *Interpretation Positive Allo-Feedback:* S believes that A's interpretation of S's previous utterance was successful.
- *Interpretation Negative Allo-Feedback:* S believes that A's interpretation of S's previous utterance was unsuccessful.
- *Perception Positive Allo-Feedback:* S believes that A's perception of S's previous utterance was successful.
- *Perception Negative Allo-Feedback:* S believes that A's perception of S's previous utterance was unsuccessful.
- *Attention Negative Allo-Feedback:* S believes that A did not pay attention to S's previous utterance.

- Feedback Elicitation Functions: S requests A to provide information about A's processing of S's previous utterance.
  - *Attention Feedback Elicitation:* S wants to know whether A is paying attention to S.
  - *Perception Feedback Elicitation:* S wants to know A's perception of S's previous utterance was successful.
  - *Interpretation Feedback Elicitation:* S wants to know A's interpretation of S's previous utterance was successful.
  - *Evaluation Feedback Elicitation:* S wants to know A's evaluation of S's previous utterance was successful.
  - *Execution Feedback Elicitation:* S wants to know A's execution of S's previous utterance was successful.

- Interaction Management Functions:
  - *Turn management functions (Turn Management Dimension)* are the communicative functions of dialogue acts which are performed in order to keep or to reallocate the speaker role. The beginning and end of a turn are associated with a reallocation of the speaker role. The units of Turn Management can thus have both a turn-unit-initial and a turn-unit-final function, which is captured by giving them a *pair* of functions, an initial and a final one.
    - *Turn-unit-initial functions:*
      - *Turn Take:* S wants to have the turn, which is available.
      - *Turn Accept:* S agrees to take the turn, which A has given to him.
      - *Turn Grab:* S wants to get the turn, which A currently has, before A assigns the turn to him or releases it.

- *Turn-unit-final functions:*
  - *Turn Keep:* S wants to keep the turn.
  - *Turn Assign:* S wants A to take the turn.
  - *Turn Release:* S wants to make the turn available to any participant.

- *Time management functions (Time Management Dimension):*
  - *Stalling:* S needs a little bit of time to formulate an utterance.
  - *Pausing:* S needs some time to do something (either in preparation of continuing the dialogue, or because something else came up which is more urgent for him to attend to) and therefore wants to suspend the dialogue for a while.

- *Contact management functions (Contact Management Dimension):*
  - *Contact Check:* S wants to establish whether A is ready to receive messages from and to send messages to S.
  - *Contact Indication:* S wants A to know that S is ready to send messages to and receive messages from A.

- *Own communication management functions: (Own Communication Management Dimension):*
  - *Self-error signal:* S wants A to know that S has made a mistake in speaking.
    - *Retraction:* S wants to withdraw something that he said within the same turn.
    - *Self-correction:* S wants to correct an error that he made within the same turn.

- *Partner communication management functions (Partner Communication Management Dimension):*
  - *Completion:* S wants to help A to complete an utterance that A is struggling to complete.
  - *Correct-misspeaking:* S wants to correct (part of) an utterance by A, believing that A made a speaking error.

- *Dialogue structure management functions (Discourse Structure Management Dimension):*
  - *Opening:* S wants A to know that S is ready and willing to engage in a dialogue with A, of which the present utterance precedes any utterance with a activity-oriented function in the current dialogue utterance precedes any utterance with a activity-oriented function in the current dialogue.
  - *Preclosing:* S plans to end the current dialogue shortly.

- *Topic introduction:* S wants to introduce the topic mentioned in the semantic content.
- *Topic shift announcement:* S wants to change the topic.
  - *Topic shift:* S wants to shift the topic to the one mentioned in the semantic content.

- *Social Obligations Management Functions: (Social Obligation Management Dimension)*
  - *Salutation*
    - *Initial greeting:* S wants A to be aware of S's presence; S is aware of A's presence; S believes that S and A are in a position to exchange messages; S puts pressure on A to acknowledge this.
    - *Return greeting:* S wants A to be aware of A's presence; S is aware of A's presence; S believes that S and A are in a position to exchange messages; S is pressured to respond to an initial greeting by A addressed to S.
  - *Self-introduction*
    - *Initial self-introduction:* S wants to make himself known to A; S puts pressure on A to acknowledge this.
    - *Return self-introduction:* S wants to make himself known to A; S is pressured to do so by an initial self-introduction by A addressed to S.
  - *Apologizing*
    - *Apology:* S wants A to know that S regrets having made an error in perceiving, understanding, evaluating, or executing an utterance by A, or not having paid attention to, perceived well, or misunderstood an utterance from A, or being unable to evaluate or execute an utterance from A; S puts pressure on A to acknowledge this.
    - *Apology-downplay:* S wants to mitigate A's feelings of regret; S has been pressured to respond to an apology by A addressed to S.
  - *Gratitude*
    - *Thanking:* S wants A to know that S is grateful for what A has done in the current dialogue; S puts pressure on A to acknowledge this.
    - *Thanking-downplay:* S wants to mitigate A's feelings of gratitude; S has been pressured to respond to a thanking act by A addressed to S.
  - *Valediction*
    - *Initial goodbye:* S wants A to know that S intends the current utterance to be his final contribution to the dialogue; S puts pressure on A to acknowledge this.
    - *Return goodbye:* S wants A to know that S intends the current utterance to be his final contribution to the dialogue; S wants to acknowledge his

awareness that A want his (A's) last utterance to be his final contribution to the dialogue; S has been pressured to respond to an initial goodbye by A addressed to S.

Emotion types

In this appendix, we provide the definitions of the 22 emotion types as they are presented in the OCC model (Ortony et al., 1988), as well as the four new emotion types we defined ourselves.

1. **JOY EMOTIONS**
   TYPE SPECIFICATION: (pleased about) a desirable event.
   TOKENS: contented, cheerful, delighted, ecstatic, elated, euphoric, feeling good, glad, happy, joyful, jubilant, pleasantly surprised, pleased, etc.
   VARIABLES AFFECTING INTENSITY:
   (1) the degree to which the event is desirable.
   EXAMPLE: The man was pleased when he realized he was to get a small inheritance from an unknown distant relative.

2. **DISTRESS EMOTIONS**
   TYPE SPECIFICATION: (displeased about) an undesirable event.
   TOKENS: depressed, distressed, displeased, dissatisfied, distraught, feeling bad, feeling uncomfortable, grief, homesick, lonely, lovesick, miserable, regret, sad, shock, uneasy, unhappy, upset, etc.
   VARIABLES AFFECTING INTENSITY:
   (1) the degree to which the event is undesirable.
   EXAMPLE: The driver was upset about running out of gas on the freeway.

3. **HAPPY-FOR EMOTIONS**
   TYPE SPECIFICATION: (pleased about) an event presumed to be desirable for someone else.
   TOKENS: delighted-for, happy-for, pleased-for, etc.
   VARIABLES AFFECTING INTENSITY:
   (1) the degree to which the desirable event for the other is desirable for oneself.
   (2) the degree to which the event is presumed to be desirable for the other person.
   (3) the degree to which the other person deserved the event.
   (4) the degree to which the other person is liked.
   EXAMPLE: Fred was happy for his friend Mary because she won a thousand dollars.

4. **SORRY-FOR EMOTIONS**
   TYPE SPECIFICATION: (displeased about) an event presumed to be undesirable for someone else.
   TOKENS: compassion, pity, sad-for, sorry-for, sympathy, etc.
   VARIABLES AFFECTING INTENSITY:
   (1) the degree to which the undesirable event for the other is undesirable for oneself.
   (2) the degree to which the event is presumed to be undesirable for the other person.
   (3) the degree to which the other did not deserve the event.
   (4) the degree to which the other person is liked.
   EXAMPLE: Fred was sorry for his friend Mary because her husband was killed in a car crash.

5. **RESENTMENT EMOTIONS**
   TYPE SPECIFICATION: (displeased about) an event presumed to be desirable for someone else.
   TOKENS: envy, jealousy, resentment, etc.
   VARIABLES AFFECTING INTENSITY:
   (1) the degree to which the desirable event for the other person is undesirable for oneself.
   (2) the degree to which the event is presumed to be desirable for the other person.
   (3) the degree to which the other person did not deserve the event.
   (4) the degree to which the other person is not liked.
   EXAMPLE: The executive resented the large pay raise awarded to a colleague whom he considered incompetent.

6. **GLOATING EMOTIONS**
   TYPE SPECIFICATION: (pleased about) an event presumed to be undesirable for someone else.
   TOKENS: gloating, Schadenfreude, etc.
   VARIABLES AFFECTING INTENSITY:
   (1) the degree to which the undesirable event for the other person is desirable for oneself.
   (2) the degree to which the event is presumed to be undesirable for the other person.
   (3) the degree to which the other person deserved the event.
   (4) the degree to which the other person is liked.
   EXAMPLE: Political opponents of Richard Nixon gloated over his ignominious departure from office.

7. **HOPE EMOTIONS**
   TYPE SPECIFICATION: (pleased about) the prospect of a desirable event.

TOKENS: anticipation, anticipatory excitement, excitement, expectancy, hope, hopeful, looking forward to, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event is desirable.
(2) the likelihood of the event.
EXAMPLE: As she thought about the possibility of being asked to the dance, the girl was filled with hope.

8. **FEAR EMOTIONS**
TYPE SPECIFICATION: (displeased about) the prospect of an undesirable event.
TOKENS: apprehensive, anxious, cowering, dread, fear, fright, nervous, petrified, scared, terrified, timid, worried, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event is undesirable.
(2) the likelihood of the event.
EXAMPLE: The employee, suspecting he was no longer needed, feared that he would be fired.

9. **SATISFACTION EMOTIONS**
TYPE SPECIFICATION: (pleased about) the confirmation of the prospect of a desirable event.
TOKENS: gratification, hopes-realized, satisfaction, etc.
VARIABLES AFFECTING INTENSITY:
(1) the intensity of the attendant Hope emotion.
(2) the effort expended in trying to attain the event.
(3) the degree to which the event is realized.
EXAMPLE: When she realized that she was indeed being asked to go to the dance by the boy of her dreams, the girl was gratified.

10. **FEARS-CONFIRMED EMOTIONS**
TYPE SPECIFICATION: (displeased about) the confirmation of the prospect of an undesirable event.
TOKENS: fears-confirmed, worst fears realized.
VARIABLES AFFECTING INTENSITY:
(1) the intensity of the attendant Fear emotion.
(2) the effort expended in trying to prevent the event.
(3) the degree to which the event is realized.
EXAMPLE: The employee's fears were confirmed when he learned that he was indeed going to be fired.

11. **RELIEF EMOTIONS**
TYPE SPECIFICATION: (pleased about) the disconfirmation of the prospect of an undesirable event.
TOKENS: relief
VARIABLES AFFECTING INTENSITY:

(1) the intensity of the attendant Fear emotion.
(2) the effort expended in trying to prevent the event
(3) the degree to which the event is realized.
EXAMPLE: The employee was relieved to learn that he was not going to be fired.

12. **DISAPPOINTMENT EMOTIONS**
TYPE SPECIFICATION: (displeased about) the disconfirmation of the prospect of a desirable event.
TOKENS: dashed-hope, despair, disappointment, frustration, heartbroken, etc.
VARIABLES AFFECTING INTENSITY:
(1) the intensity of the attendant Hope emotion.
(2) the effort expended in trying to attain the event.
(3) the degree to which the event is realized.
EXAMPLE: The girl was disappointed when she realized that she would not be asked to the dance after all.

13. **PRIDE EMOTIONS**
TYPE SPECIFICATION: (approving of) one's own praiseworthy action.
TOKENS: pride
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged praiseworthiness.
(2) the strength of the cognitive unit with the actual agent.
(3) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
EXAMPLE: The woman was proud of saving the life of a drowning child.

14. **SELF-REPROACH EMOTIONS**
TYPE SPECIFICATION: (disapproving of) one's own blameworthy action.
TOKENS: embarrassment, feeling guilty, mortified, self-blame, self-condemnation, self-reproach, shame, (psychologically) uncomfortable, uneasy, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged blameworthiness.
(2) the strength of the cognitive unit with the actual agent.
(3) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
EXAMPLE: The spy was ashamed of having betrayed his country.

15. **APPRECIATION EMOTIONS**
TYPE SPECIFICATION: (approving of) someone else's praiseworthy action.
TOKENS: admiration, appreciation, awe, esteem, respect, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged praiseworthiness.
(2) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
EXAMPLE: The physicist's colleagues admired him for his Nobel-prize-winning

work.

16. **REPROACH EMOTIONS**
TYPE SPECIFICATION: (disapproving of) someone else's blameworthy action.
TOKENS: appalled, contempt, despise, disdain, indignation, reproach, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged blameworthiness.
(2) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
EXAMPLE: Many people despised the spy for having betrayed his country.

17. **GRATITUDE EMOTIONS**
TYPE SPECIFICATION: (approving of) someone else's praiseworthy action and (being pleased about) the related desirable event.
TOKENS: appreciation, gratitude, feeling indebted, thankful, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged praiseworthiness.
(2) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
(3)the degree to which the event is desirable.
EXAMPLE: The woman was grateful to the stranger for saving the life of her child

18. **ANGER EMOTIONS**
TYPE SPECIFICATION: (disapproving of) someone else's blameworthy action and (being displeased about) the related undesirable event.
TOKENS: anger, annoyance, exasperation, fury, incensed, indignation, irritation, livid, offended, outrage, rage, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged blameworthiness.
(2) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
(3) the degree to which the event is undesirable.
EXAMPLE: The woman was angry with her husband for forgetting to buy the groceries.

19. **GRATIFICATION EMOTIONS**
TYPE SPECIFICATION: (approving of) one's own praiseworthy action and (being pleased about) the related desirable event.
TOKENS: gratification, pleased-with-oneself, self-satisfaction, smug, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged praiseworthiness.
(2) the strength of the cognitive unit with the agent.
(3) deviations of the agent's action from person/role-based expectations (i.e.

unexpectedness).
(4) the degree to which the event is desirable.
EXAMPLE: The man was gratified by his daughter's achievements.

20. **REMORSE EMOTIONS**
TYPE SPECIFICATION: (disapproving of) one's own blameworthy action and (being displeased about) the related undesirable event.
TOKENS: penitent, remorse, self-anger, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree of judged blameworthiness.
(2) the strength of the cognitive unit with the agent.
(3) deviations of the agent's action from person/role-based expectations (i.e. unexpectedness).
(4) the degree to which the event is undesirable.
EXAMPLE: The spy felt remorse at the damage he had done in betraying his country.

21. **LIKING EMOTIONS**
TYPE SPECIFICATION: (liking) an appealing object.
TOKENS: adore, affection, attracted-to, like, love, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the object is appealing.
(2) the degree of familiarity with the object.
EXAMPLE: Mary was filled with affection as she gazed at her newborn infant.

22. **DISLIKING EMOTIONS**
TYPE SPECIFICATION: (disliking) an unappealing object.
TOKENS: aversion, detest, disgust, dislike, hate, loath, repelled-by, revulsion, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the object is unappealing.
(2) the degree of familiarity with the object.
EXAMPLE: John disliked the concert so much that he left in the middle.

23. **SYMPATHY EMOTIONS**
TYPE SPECIFICATION: experiencing someone else's feeling that is unfamiliar
TOKENS: compassion, sensitivity, affinity, commiseration, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which someone else's feelings are experienced.
(2) the degree to which the agent want to alter those feelings
EXAMPLE: The girl felt sympathy for the man's feelings and although she didn't understand them, she wanted to alleviate them.

24. **EMPATHY EMOTIONS**
TYPE SPECIFICATION: understanding someone else's feeling that is familiar

TOKENS: insight, understanding, appreciation, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which someone else's feelings are understood.
EXAMPLE: the girl understood that the man was feeling hurt, but didn't want to interfere.

25. **SURPRISE EMOTIONS**
TYPE SPECIFICATION: understanding an event or action that is unexpected
TOKENS: awe, amazement, revelation, shock, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event or action is unexpected.
EXAMPLE: the man was surprise when he saw who had come to his birthday.

26. **CONFUSION EMOTIONS**
TYPE SPECIFICATION: not understanding an event or action that is unexpected
TOKENS: bewilderment, disorientation, perplexity, puzzlement, etc.
VARIABLES AFFECTING INTENSITY:
(1) the degree to which the event or action is unexpected.
(2) the degree to which the even or action is understood.
EXAMPLE:the man was confused about the presence of all the people.

Questionnaire

In this appendix, we present the questionnaire used to gather data concerning the properties of the conversational behaviors performed in the Pallium video and the participants' assumptions about the underlying internal state features of both interlocutors.

THE QUESTIONNAIRE:

Dear participant,

Thank you for filling in this questionnaire. Your participation is anonymous and voluntary. In a moment you will get to see fragments from a short video. You can repeat the fragment by pressing the play button again. After each fragment you are asked to answer a set of questions about the persons in the video. Please take a moment to consider each question and your answer carefully. You can give the answers in your own words, either in English or in Dutch. Note that there are no right or wrong answers to the questions; we are interested in what you are thinking. You can take as long as you want to complete the questions. Here are some questions that you may be asked to answer.

1. What do you think the doctor/patient wants to achieve with his behavior?
2. What do you anticipate to be the effect of the doctor's/patient's behavior on the patient/doctor and why?
3. Please describe what you think the patient/doctor is thinking and feeling at this moment.

After filling in the questions, you can continue to the next fragment by pressing the "next" button. It is not possible to go back to a previous fragment, so please make sure you have answered all questions before continuing to the next video fragment.A few suggestions for things to pay attention to while watching the fragments are:

- What is the speaker saying?
- What else is conveyed in his speech (by tone of voice, for instance)?
- What is expressed through body language (facial expressions, gestures, distance between the persons, touch, etcetera).

Doctor1 (D1): *I'm about to visit John Filpot. He is recovering from surgery. He had a bowel tumor removed and evidence of metastases. This means that John's cancer is incurable. Now, I know the surgeon has talked with John, but I don't know if John really understands his condition.*
Q1: Please describe what you think the doctor is thinking and feeling at this moment.

Patient1 (P1): *I'm feeling pretty good. Yeah I got some pain, but the surgeon got my tumor. It was attached to my bowel. But the most important thing is that the tumor is gone. I'll probably be out of here in a few days, once they can remove these tubes and get me on some decent food. I'm sure everything is going to be okay. No biggy, as my son says.*
Q2: Please describe what you think the patient is thinking and feeling at this moment.

D2: *Hey John.*
Q3: Please describe what you think the doctor is thinking and feeling at this moment.
Q4: What do you think the doctor wants to achieve with his behavior?
Q5: What do you anticipate to be the effect of the doctor's behavior on the patient?

P2: *Hello doctor Rupert.*
Q6: What do you think the patient wants to achieve with his behavior?

D3: *I'm sorry to hear the news. We were hoping that surgery and a course of chemo would cure this tumor.*
Q7: What do you think the doctor wants to achieve with his behavior?
Q8: What do you anticipate to be the effect of the doctor's behavior on the patient?

P3: *What do you mean?*
Q9: What do you think the patient wants to achieve with his behavior?
Q10: What effect do you think the doctor's previous behavior had on the patient's behavior in this fragment?
Q11: Please describe what you think the patient is thinking and feeling at this moment.

D4: *When a tumor has spread through the abdomen, like yours has, there is a limited amount we can do for you.*
Q12: What do you think the doctor wants to achieve with his behavior?
Q13: What do you anticipate to be the effect of the doctor's behavior on the patient?

P4: *What do you mean, spread through the abdomen like mine has? The surgeon said he had removed the tumor.*
Q14: What do you think the patient wants to achieve with his behavior?
Q15: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q16: Please describe what you think the patient is thinking and feeling at this moment.

D5: *Yes, The primary one. But he couldn't get all the places it had invaded; (camera shifts) we need to give you chemo for that.*
Q17: What do you think the doctor wants to achieve with his behavior?
Q18: How do you think the doctor's behavior will have an effect on the patient?
Q19: Please describe what you think the doctor is thinking and feeling at this moment.

P5: *What do you mean invaded? You keep talking about this chemo-stuff.*
Q20: What do you think the patient wants to achieve with his behavior?
Q21: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q22: Please describe what you think the patient is thinking and feeling at this moment.

D5: *Didn't the surgeon talk with you?*
Q23: What do you think the doctor wants to achieve with his behavior?
Q24: What do you anticipate to be the effect of the doctor's behavior on the patient?
Q25: Please describe what you think the doctor is thinking and feeling at this moment.

P6: *Yeah but he didn't say anything about chemo or invasive stuff. What the hell is going on here? I thought I was cured, (camera shift) but now it sounds like I'm not.*
Q26: What do you think the patient wants to achieve with his behavior?
Q27: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q28: Please describe what you think the patient is thinking and feeling at this moment.

D6: *I'm sorry John. I thought you knew.*
Q29: What do you think the doctor wants to achieve with his behavior?
Q30: What do you anticipate to be the effect of the doctor's behavior on the patient?
Q31: Please describe what you think the doctor is thinking and feeling at this moment.

P7: *Well I didn't! I don't believe it. (Pause) They have shown you the wrong chart! I'm going to be fine (camera shift). Just need a little medicine and I'll be right as rain.*
Q32: What do you think the patient wants to achieve with his behavior?
Q33: What do you anticipate to be the effect of the patient's behavior on the doctor?
Q34: Please describe what you think the patient is thinking and feeling at this moment.

D7: *Okay, John. I'll be back later.*
Q35: What do you think the doctor wants to achieve with his behavior?
Q36: Please describe what you think the doctor is thinking and feeling at this moment.

Are there any other questions which you think might be important for this ques-

tionnaire, or do you have remarks about the interaction between the doctor and the patient as a whole, or any other suggestions, please write them down here.

Thank you for participating in this questionnaire.

Bibliography

Adam, C., Gaudou, B., Herzig, A., and Longin, D. (2006). Occ's emotions: A formalization in a bdi logic. *Artificial Intelligence: Methodology, Systems, and Applications*, 12:24–32.

Allen, J. and Core, M. (1997). Damsl: Dialog act markup in several layers. http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/RevisedManual.html.

Allwood, J. (2000). An activity based approach to pragmatics. In *Abduction, belief and context in dialogue: Studies in computational pragmatics*, pages 47–80. Amsterdam, John Benjamins.

Allwood, J. S. (1976). *Linguistic communication as action and cooperation: A study in pragmatics*. Dept. of Linguistics, University of Göteborg, Göteborg.

Austin, J. L. (1962). *How to do things with words*. Harvard University Press, Cambridge, MA.

Bach, J. (2003). The micropsi agent architecture. In *Proceedings of ICCM-5, International Conference on Cognitive Modeling*.

Baile, W., Buckman, R., Lenzi, R., Glober, G., Beale, E., and Kudelka, A. (2000). Spikes-a six-step protocol for delivering bad news: application to the patient with cancer. *Oncologist*, 5(4):302–311.

Bickmore, T. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Bickmore, T. (2008). Framing and interpersonal stance in relational agents. In *Why Conversational Agents do what they do. Functional Representations for Generating Conversational Agent Behavior*.

Bickmore, T. W. and Sidner, C. L. (2006). Towards plan-based health behavior change counseling systems. In *Proceedings of AAAI Spring Symposium on Argumentation for Consumers of Healthcare*.

Bloomfield, L. (1933). *Language*. Holt publishing, New York, NY.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press, Cambridge, MA.

Buckman, R. and Kason, Y. (1992). *How to break bad news: a guide for health care professionals*. How to Break Bad News. Johns Hopkins University Press.

Bui, T. H. (2006). Multimodal dialogue management - state of the art. Technical Report TR-CTIT-06-01, Centre for Telematics and Information Technology, University of Twente, Enschede.

Bulman, R. and Wortman, C. (1977). *Attributions of Blame and Coping in the Real World: Severe Accident Victims React to Their Lot*. National Emergency Training Center.

Bunt, H. (1994). Context and dialogue control. *THINK Quarterly*, 3:19–31.

Bunt, H. (2005). A framework for dialogue act specification. In *Presented at the Fourth ISO/SIGSEM Workshop on Multimodal Semantic Representation*, Tilburg, The Netherlands.

Bunt, H. (2006). Dimensions in dialogue act annotation. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, volume 6, pages 919–924.

Bunt, H. (2009). The dit++ taxonomy for functional dialogue markup. In Dirk Heylen, Catherine Pelachaud, R. C. and Traum, D., editors, *Proceedings of the AAMAS 2009 Workshop "Towards a Standard Markup Language for Embodied Dialogue Acts" (EDAML 2009)*.

Bunt, H. (2011). The semantics of dialogue acts. In *Proceedings of the Ninth International Conference on Computational Semantics*, IWCS '11, pages 1–13, Stroudsburg, PA, USA. Association for Computational Linguistics.

Bunt, H., Alexandersson, J., Choe, J.-W., Fang, A. C., Hasida, K., Petukhova, V., Popescu-Belis, A., and Traum, D. (2012). Iso 24617-2: A semantically-based standard for dialogue annotation. In Calzolari, N., Choukri, K., Declerck, T., Doan, M. U., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., and Piperidis, S., editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey. European Language Resources Association (ELRA).

Bunt, H. and Black, W. (2000). Dialogue pragmatics and context specification. In *Abduction, Belief and Context in Dialogue; studies in computational*, pages 81–150. John Benjamins.

Carletta, J., Isard, A., Isard, S., Kowtko, J. C., Doherty-Sneddon, G., and Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13–31.

Carver, C., Scheier, M., and Weintraub, J. (1989). Assessing coping strategies: a theoretically based approach. *Journal of Personality and Social Psychology*, 56(2):267–283.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., and Yan, H. (1999). Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 520–527, New York, NY, USA. ACM.

Clark, H. H. (1996). *Using Language*. Cambridge University Press, Cambridge, MA.

Cohen, F. and Lazarus, R. (1973). Active coping processes, coping dispositions, and recovery from surgery. *Psychosomatic Medicine*, 35(5):375–389.

Cohen, P. (1997). Dialogue modeling. In Cole, R., editor, *Survey of the state of the art in human language technology*, pages 204–210. Cambridge University Press, New York, NY, USA.

Cohen, P. R. and Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261.

Copp, G. (1998). A review of current theories of death and dying. *Journal of Advanced Nursing*, 28(2):382–390.

Corr, C. (1993). Coping with dying: Lessons that we should and should not learn from the work of elisabeth kübler-ross. *Death Studies*, 17(1):69–83.

Dignum, F., Kinny, D., and Sonenberg, L. (2002). Motivational attitudes of agents: On desires, obligations, and norms. In *Revised Papers from the Second International Workshop of Central and Eastern Europe on Multi-Agent Systems: From Theory to Practice in Multi-Agent Systems*, CEEMAS '01, pages 83–92, London, UK, UK. Springer-Verlag.

Eggly, S., Penner, L., Albrecht, T., Cline, R., Foster, T., Naughton, M., Peterson, A., and Ruckdeschel, J. (2006). Discussing bad news in the outpatient oncology clinic: rethinking current communication guidelines. *Journal of Clinical Oncology*, 24(4):716–719.

Ekman, P. (1971). *Universals and Cultural Differences in Facial Expressions of Emotion*. University of Nebraska Press.

Elliot, C. (1992). *The Affective Reasoner: A Process model of emotions in a multi-agent system*. PhD thesis, Northwestern University.

Falcone, R. and Castelfranchi, C. (2001). Trust and deception in virtual societies. In Castelfranchi, C. and Tan, Y.-H., editors, *Trust and Deception in Virtual Societies*, chapter Social Trust: A Cognitive Approach, pages 55–90. Kluwer Academic Publishers, Norwell, MA, USA.

Fallowfield, L. and Jenkins, V. (2004). Communicating sad, bad, and difficult news in medicine. *Lancet*, 363(9405):312–319.

Folkman, S. and Lazarus, R. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, 21(3):219–239.

Folkman, S. and Lazarus, R. (1985). If it changes it must be a process: study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, 48(1):150–170.

Frege, G. (1948 / 1892). On sense and reference. *Philosophical Review*, 57:209–230.

Gillotti, C., Thompson, T., and McNeilis, K. (2002). Communicative competence in the delivery of bad news. *Social Science and Medicine*, 54(7):1011–1023.

Goldsmith, J., Wittenberg-Lyles, E., Villagran, M., and Sanchez-Reilly, S. (2008). Inefficacy of the spikes strategy in the terminal bad news context: Critical conversations of medical students, standardized patients and faculty. Paper presented at the annual meeting of the NCA 94th Annual Convention.

Goodwin, C. (1981). *Conversational Organization: Interaction Between Speakers and Hearers*. Academic Press, New York, NY.

Gratch, J. (2000). émile: Marshalling passions in training and education. In *Proceedings of the fourth international conference on Autonomous agents*, AGENTS '00, pages 325–332, New York, NY, USA. ACM.

Gratch, J., Mao, W., and Marsella, S. (2006a). Modeling social emotions and social attributions. In *Cognitive Modeling and Multi-Agent Interaction, Run Sun (ed.), Cambridge University Press, 2006, pp 219-251.*

Gratch, J. and Marsella, S. C. (2001). Modeling emotions in the mission rehearsal exercise. In *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*, pages 457–466, Orlando, FL.

Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S. C., Morales, M., van der Werf, R. J., and Morency, L.-P. (2006b). Virtual rapport. In *Lecture Notes in Computer Science*, volume 4311, pages 14–27, Marina del Rey, CA.

Heuvelink, A., van den Bosch, K., van Doesburg, W., and Harbers, M. (2009). Intelligent agent supported training in virtual simulations. In *Proceedings of the NATO HFM-169 Workshop on Human Dimensions in Embedded Virtual Simulation, , NATO Human Factors and Medicine Panel*. AAAI Press.

Hospers, M., Kroezen, E., Nijholt, A., op den Akker, H., and Heylen, D. (2003). An agent-based intelligent tutoring system for nurse education. In Nealon, J. and Moreno, A., editors, *Ap-*

plications of Software Agent Technology in the Health Care Domain, Whitestein Series in Software Agent Technologies, pages 143–159. Birkhauser Publishing Ltd, Basel, Switzerland. ISBN=978-3-7643-2662-3.

Jacobs, S. (1993). *Pathologic grief: maladaption to loss*. American Psychiatric Press.

Jurafsky, D. and Martin, J. H. (2000). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition.

Kastenbaum, R. and Costa jr., P. (1977). Psychological perspectives on death. *Annual review of psychology*, 28(1):225–249.

Kopp, S., Gesellensetter, L., Krämer, N. C., and Wachsmuth, I. (2005). A conversational agent as museum guide - design and evaluation of a real-world application. In Panayiotopoulos, T., editor, *Intelligent virtual agents: Proceedings 5th International Conference*, pages 329–343. Springer.

Kopp, S., Jung, B., Leßmann, N., and Wachsmuth, I. (2003). Max - a multimodal assistant in virtual reality construction. *KI - Künstliche Intelligenz*, 4(03):11–17.

Kübler-Ross, E. (1969). *On death and dying*. Macmillan, New York.

Lazarus, R. (1999). *Stress and Emotion: A New Synthesis*. Free Association Books.

Lazarus, R. S. and Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Publishing Company, New York.

Leventhal, H. and Scherer, K. (1987). The Relationship of Emotion to Cognition: A Functional Approach to a Semantic Controversy. *Cognition & Emotion*, 1(1):3–28.

Lyons, J. (1969). *Introduction to Theoretical Linguistics*. Cambridge University Press, Cambridge.

Maciejewski, P., Zhang, B., Block, S., and Prigerson, H. (2007). An empirical examination of the stage theory of grief. *JAMA: The Journal of the American Medical Association*, 297(7):716–723.

Marsella, S. and Gratch, J. (2006). Ema: A computational model of appraisal dynamics. In *European Meeting on Cybernetics and Systems Research*.

Marsella, S. C. and Gratch, J. (2009). EMA: A process model of appraisal dynamics. *Journal of Cognitive Systems Research*, 10(1):70–90.

Miranda, J. and Brody, R. (1992). Communicating bad news. *Western Journal of Medicine*, 156(1):83–85.

Ochs, M., Niewiadomski, R., Pelachaud, C., and Sadek, D. (2005). Intelligent expressions of emotions. In *Affective Compunting and Intelligent Interaction*, pages 707–714. Springer.

Orlander, J., Fincke, B., Hermanns, D., and Johnson, G. (2002). Medical residents' first clearly remembered experiences of giving bad news. *Journal of General Internal Medicine*, 17(11):825–831.

Ortony, A., Clore, G. L., and Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press.

Pearlin, L. and Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, 19(1):2–21.

Pereira, D., Oliveira, E., and Moreira, N. (2008). Formal modelling of emotions in bdi agents. In Sadri, F. and Satoh, K., editors, *Computational Logic in Multi-Agent Systems*, volume 5056 of *Lecture Notes in Computer Science*, pages 62–81. Springer Berlin Heidelberg.

Poggi, I., Pelachaud, C., de Rosis, F., Carofiglio, V., and de Carolis, B. (2005). GRETA. A Believable Embodied Conversational Agent. In Paggio, P. and Jongejan, B., editors, *Multimodal Communication in Virtual Environments*, pages 27–45. Springer Netherlands.

Ptacek, J. and Eberhardt, T. L. (1996). Breaking bad news: A review of the literature. *JAMA: The Journal of the American Medical Association*, 276(6):496–502.

Rao, A. S. and Georgeff, M. P. (1991). Modeling rational agents within a bdi-architecture. In Allen, J. F., Fikes, R., and Sandewall, E., editors, *KR*, pages 473–484. Morgan Kaufmann.

Rickel, J. and Johnson, W. L. (1997). Steve: An animated pedagogical agent for procedural training in virtual environments (extended abstract). *SIGART Bulletin*, 8:16–21.

Rosenbaum, M., Ferguson, K., and Lobas, J. (2004). Teaching medical students and residents skills for delivering bad news: a review of strategies. *Academic Medicine*, 79(2):107–117.

Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4, Part 1):696–735.

Sadek, M. (1991). Dialogue acts are rational plans. In *Proceedings of the ESCA/ETR workshop on multi-modal dialogue.*

Scheier, M. and Carver, C. (1988). A model of behavioral self-regulation: Translating intention into action. *Advances in experimental social psychology*, 21:303–346.

Scherer, K. (2005). What are emotions? and how can they be measured? *Social Science Information*, 44:695–729.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In Ekman, P. & Scherer, K. R., editor, *Approaches to emotion*, chapter 14, pages 293–317. NJ: Erlbaum, Hillsdale, k.r. scherer & p. ekman (eds.) edition.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge.

Searle, J. R. (1975). Indirect speech acts. In Cole, P. and Morgan, J., editors, *Syntax and Semantics 3: Speech Acts*, pages 59–82. Academic Press, New York.

Searle, J. R. (1989). How performatives work. *Linguistics and Philosophy*, 12(5):535–558.

Steunebrink, B. R. (2010). *The logical structure of emotions*. PhD thesis, Utrecht University.

ter Maat, M. (2011). *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent*. Phd thesis, University of Twente. publisher: Twente University Press, publisherlocation: Enschede, ISBN: 978-94-6191-105-6, Numberofpages: 119.

ter Maat, M. and Heylen, D. (2011). Flipper: An information state component for spoken dialogue systems. In Vilhjálmsson, H., Kopp, S., Marsella, S., and Thórisson, K., editors, *Proceedings of the 10th international conference on Intelligent Virtual Agents (IVA 2011)*, volume 6895 of *Lecture Notes in Computer Science*, pages 470–472, Berlin. Springer Verlag. ISBN=978-3-642-23973-1, ISSN=0302-9743.

Theune, M., Hofs, D., and van Kessel, M. (2007). The virtual guide: A direction giving embodied conversational agent. In *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pages 2197–2200, Bonn. International Speech Communication Association (ISCA). ISSN=1990-9772.

Thórisson, K. R. (1997). Gandalf: An embodied humanoid capable of real-time multimodal dialogue with people. In *Agents*, pages 536–537.

Thórisson, K. R. (2002). Natural turn-taking needs no manual: Computational theory and model, from perception to action.

Traum, D. (2000). 20 questions for dialogue act taxonomies. *Journal of Semantics*, 17(1):7–30.

Traum, D. and Larsson, S. (2003). The information state approach to dialogue management. In *Current and New Directions in Discourse and Dialogue*, pages 325–353. Springer.

Traum, D. R. (1993). Mental state in the trains-92 dialogue manager. In *In Working Notes of the AAAI Spring Symposium on Reasoning about Mental States: Formal Theories and Applications*, pages 143–149. AAAI Press.

Traum, D. R. and Hinkelman, E. A. (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8:575–599.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236):433–460.

van Welbergen, H., Reidsma, D., Ruttkay, Z., and Zwiers, J. (2010). Elckerlyc - a bml realizer for continuous, multimodal interaction with a virtual human. *Journal on Multimodal User Interfaces*, 3(4):271–284. impact=0.000, ISSN=1783-7677.

Weisman, Avery D.; Worden, J. W. (1976). The existential plight in cancer: Significance of the first 100 days. *International Journal of Psychiatry in Medicine*, 7(1):1–15.

Wooldridge, M. (2000). *Reasoning about rational agents*. MIT press.

Summary

One of the most characteristic aspects of natural communication is a human's ability to make assumptions about the internal mental state of his or her interlocutor. This enables the speaker to do two important things. First, it allows him to form beliefs about the interlocutor's thoughts and feelings that might underlie his conversational behaviors. Subsequently, it enables the speaker to select and perform appropriate response behaviors that might alter the interlocutor's thoughts and / or feelings, and consequently result in desired conversation behaviors. This in turn will lead to a desired conversation. I argue that the thoughts and feelings (i.e. the features of his internal state) of the speaker as well as his beliefs about the internal state features of his interlocutor are used in the behavior processing and behavior selection processes.

The research presented in this thesis focuses on the way these conversational mechanisms operate. The goals of the research were to gain more insight in how the cognitive processes involved in processing perceived conversational behaviors and in selecting appropriate response behaviors operate. In addition, I aimed to determine how these cognitive processes and the thoughts and feelings associated with them might be represented. In order to restrict the scope of the research, I focus my studies on the conversations conducted in medical bad news situations.

In order to achieve the goals of the research, several actions were performed. First, several linguistic and psychological theories and models that focus on various aspects of dialogues, such as Speech Act Theory and Dialogue Act Theory, were studied. This was done to see how the meaning and function of utterances performed in conversational behaviors could be determined. In addition, various theories and models that describe (parts of) the cognitive processes and internal state features involved in dialogues were studied.

Next, I looked into various models that describe how bad news conversations are, and should be, conducted, from the perspective of both the provider of the bad news as well as the receiver. Particular consideration was given to the use and workings of coping mechanisms, such as stage theories of grief.

After studying the various theories and models, two analyses of the conversational behaviors performed in a simulated bad news conversation were conducted. For the first analysis, I made a detailed annotation of the properties of the conversational behaviors and the displays of various internal state features observed. The second analyses consisted of an online questionnaire concerning the same conversational behaviors. The aim of the questionnaire was to gain insight in how conversational behaviors are perceived and which associated internal state features can be ascribed to the speaker.

Based on findings of both the literature study and the data analyses, a cognitive dialogue model was constructed, in which the various internal state features and processes involved in processing and selecting conversational behaviors were represented. Effort was made to keep the representation of the processes as natural and human-like as possible. The idea behind this is that by enabling the processes (and the associated internal state features) to be recognized

intuitively, it is easier to comprehend why the cognitive dialogue model selects a particular response behavior, when it provides feedback to an interlocutor. Such intuitive understanding of the feedback might prove useful in a dialogue system in which the cognitive dialogue model is used to model the role of a participant in a conversational situation, such as practicing bad news situations. I conclude my thesis with some of the benefits and shortcomings of my cognitive dialogue model and provide several suggestions on how to extend and improve my research through future work.

Samenvatting

Eén van de meest karakteristieke aspecten van natuurlijke communicatie is dat mensen in staat zijn om aannames te maken over de interne, mentale toestand van zijn of haar gesprekspartner. Dit stelt een persoon in staat om twee belangrijke dingen te doen tijdens een gesprek. Ten eerste geeft het hem de mogelijkheid om een mentaal beeld te vormen over de gedachtes en gevoelens van zijn gesprekspartner die ten grondslag zouden kunnen liggen aan diens conversationele gedrag. Daarnaast zorgt het ervoor dat de persoon gepaste conversationele reactieve gedragingen kan selecteren en uiten, die mogelijk de gedachtes en/of gevoelens van de gesprekspartner kunnen aanpassen. Het beoogde gevolg hiervan is dat de gesprekspartner dan conversationeel gedragingen gaat vertonen welke gewenst zijn door de persoon. Ik betoog dat de gedachtes en gevoelens (dat wil zeggen, de kenmerken van de interne toestand) van de persoon, samen met het mentale beeld dat hij heeft gevormd over de interne toestand van zijn gesprekspartner, gebruikt worden in de cognitieve processen die conversationeel gedragingen verwerken en selecteren.

Het onderzoek dat in deze thesis gepresenteerd wordt richt zich op de manier waarop diverse conversationele mechanismen in elkaar steken. Het doel van dit onderzoek was om meer inzicht te verschaffen in de werking van de cognitieve processen die betrokken zijn bij het verwerken van waargenomen conversationele gedragingen en bij het selecteren van gepaste conversationele reacties. Daarna heb ik getracht om vast te stellen hoe deze cognitieve processen en de bijbehorende gedachtes en gevoelens gerepresenteerd kunnen worden. Om de reikwijdte van het onderzoek te beperken heb ik mij gefocused op gesprekken die worden gehouden in medische slecht nieuws situaties.

Om de onderzoeksdoelen te bereiken zijn verschillende stappen ondernomen. Eerst zijn verscheidene linguistische en psychologische theorieën en modellen (zoals Speech Act Theory en Dialogue Act Theory) bestudeerd, welke zich focussen op verschillende aspecten van dialogen. Het doel hiervan was om uit te zoeken hoe de betekenis en functie van taaluitingen in conversationele gedragingen kan worden bepaald. Daarnaast is er gekeken naar verscheidene theorieën en modellen die een beschrijving geven van (delen van) de cognitieve processen en de kenmerken van de interne toestanden die betrokken zijn in dialogen.

Vervolgens heb ik vescheidene modellen bestudeerd die beschrijven hoe slecht nieuws gesprekken worden, en zouden moeten worden gevoerd, vanuit het perspectief van zowel de brenger als de ontvanger. Speciale aandacht is besteed aan het gebruik en de werking van coping mechanismes zoals Stage Theories of Grief.

Na het bestuderen van de verschillende theorieën en modellen zijn twee analyses uitgevoerd van de conversationele gedragingen in een gesimuleerd slecht nieuws gesprek. De eerste analyse is gedaan aan de hand van een gedetaileerde annotatie die ik gemaakt heb van de geobserveerde eigenschappen van conversationeel gedragingen en de weergave van verschillende interne toestand kenmerken daarin. De tweede analyse is gebaseerd op de antwoorden van een online questionnaire betreffende dezelfde conversationele gedragingen. Het doel van de questionnaire was inzicht te krijgen in hoe conversationele gedragingen worden

waargenomen en welke geassocieerde kenmerken van de interne toestand kunnen worden toegeschreven aan de spreker.

Op basis van de bevindingen van zowel de literatuurstudie als de twee uitgevoerde analyses is een cognitief dialoog model ontwikkeld waarin de verschillende kenmerken van een interne toestand alsmede de verschillende processen betrokken bij het verwerken en selecteren van conversationeel gedrag zijn gerepresenteerd. Bij de ontwikkeling van het dialoog model is er moeite gedaan om de representatie van de kenmerken en de processen zo natuurlijk en mensachtig mogelijk te houden. Het idee hierachter is dat door te zorgen dat de cognitieve processen (en de verwante kenmerken van de interne toestand) intuitief herkend kunnen worden, het gemakkelijker is om te begrijpen waarom het cognitieve dialoog model een specifiek conversationeel gedrag zal selecteren als reactie op het gedrag van de gesprekspartner. De mogelijkheid om intuitief te begrijpen waarom de reactieve feedback wordt gegeven, kan nuttig zijn in een dialoog systeem waarin het cognitieve dialoog model gebruikt wordt om de rol van gesprekspartner te modeleren, zoals bijvoorbeeld in een virtuele patient in een slecht nieuws gesprek simulator.

Als conclusie betoog ik dat door de reikwijdte van het ontwikkelde cognitief dialoog model een breder en meer geïnformeerd beeld gevormd kan worden over de werking van de cognitieve processen en de verwante kenmerken van de interne toestand betrokken bij conversaties, en over de complexe verbanden die er tussen de processen en kenmerken bestaan. Volgens mij kan deze informatie ons helpen om dialoog systemen te ontwikkelen die natuurlijker en mensachtiger conversationeel gedrag kunnen vertonen.

SIKS dissertation series

Since 1998, all dissertations written by Ph.D.-students who have conducted their research under auspices of a senior research fellow of the SIKS research school are published in the SIKS Dissertation Series.

2015-13 Giuseppe Procaccianti(VU), *Energy-Efficient Software.*
2015-12 Julie M. Birkholz (VU), *Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks.*
2015-11 Yongming Luo(TUE), *Designing algorithms for big graph datasets: A study of computing bisimulation and joins.*
2015-10 Henry Hermans (OUN), *OpenU: design of an integrated system to support lifelong learning.*
2015-09 Randy Klaassen(UT), *HCI Perspectives on Behavior Change Support Systems.*
2015-08 Jie Jiang (TUD), *Organizational Compliance: An agent-based model for designing and evaluating organizational interactions.*
2015-07 Maria-Hendrike Peetz (UvA), *Time-Aware Online Reputation Analysis.*
2015-06 Farideh Heidari (TUD), *Business Process Quality Computation - Computing Non-Functional Requirements to Improve Business Processes.*
2015-05 Christoph Bösch (UT), *Cryptographically Enforced Search Pattern Hiding.*
2015-04 Howard Spoelstra (OUN), *Collaborations in Open Learning Environments.*
2015-03 Twan van Laarhoven (RUN), *Machine learning for network data.*
2015-02 Faiza Bukhsh (UvT), *Smart auditing: Innovative Compliance Checking in Customs Controls.*
2015-01 Niels Netten (UvA), *Machine Learning for Relevance of Information in Crisis Response.*
2014-47 Shangsong Liang (UVA), *Fusion and Diversification in Information Retrieval.*
2014-46 Ke Tao (TUD), *Social Web Data Analytics: Relevance, Redundancy, Diversity.*
2014-45 Birgit Schmitz (OUN), *Mobile Games for Learning: A Pattern-Based Approach.*
2014-44 Paulien Meesters (UvT), *Intelligent Blauw. Met als ondertitel: Intelligence-gestuurde politiezorg in gebiedsgebonden eenheden.*
2014-43 Kevin Vlaanderen (UU), *Supporting Process Improvement using Method Increments.*
2014-42 Carsten Eijckhof (CWI/TUD), *Contextual Multidimensional Relevance Models.*
2014-41 Frederic Hogenboom (EUR), *Automated Detection of Financial Events in News Text.*
2014-40 Walter Omona (RUN), *A Framework for Knowledge Management Using ICT in Higher Education.*
2014-39 Jasmina Maric (UvT), *Web Communities, Immigration, and Social Capital.*
2014-38 Danny Plass-Oude Bos (UT), *Making brain-computer interfaces better: improving usability through post-processing.*
2014-37 Maral Dadvar (UT), *Experts and Machines United Against Cyberbullying.*
2014-36 Joos Buijs (TUE), *Flexible Evolutionary Algorithms for Mining Structured Process Mod-*

els.
2014-35 Joost van Ooijen (UU), *Cognitive Agents in Virtual Worlds: A Middleware Design Approach.*
2014-34 Christina Manteli (VU), *The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems.*
2014-33 Tesfa Tegegne (RUN), *Service Discovery in eHealth.*
2014-32 Naser Ayat (UvA), *On Entity Resolution in Probabilistic Data.*
2014-31 Leo van Moergestel (UU), *Agent Technology in Agile Multiparallel Manufacturing and Product Support .*
2014-30 Peter de Cock (UvT), *Anticipating Criminal Behaviour.*
2014-29 Jaap Kabbedijk (UU), *Variability in Multi-Tenant Enterprise Software .*
2014-28 Anna Chmielowiec (VU), *Decentralized k-Clique Matching.*
2014-27 Rui Jorge Almeida (EUR), *Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty.*
2014-26 Tim Baarslag (TUD), *What to Bid and When to Stop.*
2014-25 Martijn Lappenschaar (RUN), *New network models for the analysis of disease interaction.*
2014-24 Davide Ceolin (VU), *Trusting Semi-structured Web Data.*
2014-23 Eleftherios Sidirourgos (UvA/CWI), *Space Efficient Indexes for the Big Data Era.*
2014-22 Marieke Peeters (UU), *Personalized Educational Games - Developing agent-supported scenario-based training.*
2014-21 Kassidy Clark (TUD), *Negotiation and Monitoring in Open Environments.*
2014-20 Mena Habib (UT), *Named Entity Extraction and Disambiguation for Informal Text: The Missing Link.*
2014-19 Vinicius Ramos (TUE), *Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support.*
2014-18 Mattijs Ghijsen (VU), *Methods and Models for the Design and Study of Dynamic Agent Organizations.*
2014-17 Kathrin Dentler (VU), *Computing healthcare quality indicators automatically: Secondary Use of Patient Data and Semantic Interoperability.*
2014-16 Krystyna Milian (VU), *Supporting trial recruitment and design by automatically interpreting eligibility criteria.*
2014-15 Natalya Mogles (VU), *Agent-Based Analysis and Support of Human Functioning in Complex Socio-Technical Systems: Applications in Safety and Healthcare.*
2014-14 Yangyang Shi (TUD), *Language Models With Meta-information.*
2014-13 Arlette van Wissen (VU), *Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains.*
2014-12 Willem van Willigen (VU), *Look Ma, No Hands: Aspects of Autonomous Vehicle Control.*
2014-11 Janneke van der Zwaan (TUD), *An Empathic Virtual Buddy for Social Support.*
2014-10 Ivan Salvador Razo Zapata (VU), *Service Value Networks.*
2014-09 Philip Jackson (UvT), *Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language.*
2014-08 Samur Araujo (TUD), *Data Integration over Distributed and Heterogeneous Data Endpoints.*
2014-07 Arya Adriansyah (TUE), *Aligning Observed and Modeled Behavior.*
2014-06 Damian Tamburri (VU), *Supporting Networked Software Development.*
2014-05 Jurriaan van Reijsen (UU), *Knowledge Perspectives on Advancing Dynamic Capability.*
2014-04 Hanna Jochmann-Mannak (UT), *Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation.*
2014-03 Sergio Raul Duarte Torres (UT), *Information Retrieval for Children: Search Behavior*

and Solutions.
2014-02 Fiona Tuliyano (RUN), *Combining System Dynamics with a Domain Modeling Method.*
2014-01 Nicola Barile (UU), *Studies in Learning Monotone Models from Data.*
2013-43 Marc Bron (UVA), *Exploration and Contextualization through Interaction and Concepts.*
2013-42 Léon Planken (TUD), *Algorithms for Simple Temporal Reasoning.*
2013-41 Jochem Liem (UVA), *Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning.*
2013-40 Pim Nijssen (UM), *Monte-Carlo Tree Search for Multi-Player Games.*
2013-39 Joop de Jong (TUD), *A Method for Enterprise Ontology based Design of Enterprise Information Systems.*
2013-38 Eelco den Heijer (VU), *Autonomous Evolutionary Art.*
2013-37 Dirk Börner (OUN), *Ambient Learning Displays.*
2013-36 Than Lam Hoang (TUe), *Pattern Mining in Data Streams.*
2013-35 Abdallah El Ali (UvA), *Minimal Mobile Human Computer Interaction.*
2013-34 Kien Tjin-Kam-Jet (UT), *Distributed Deep Web Search.*
2013-33 Qi Gao (TUD), *User Modeling and Personalization in the Microblogging Sphere.*
2013-32 Kamakshi Rajagopal (OUN), *Networking For Learning; The role of Networking in a Lifelong Learner's Professional Development.*
2013-31 Dinh Khoa Nguyen (UvT), *Blueprint Model and Language for Engineering Cloud Applications.*
2013-30 Joyce Nakatumba (TUE), *Resource-Aware Business Process Management: Analysis and Support.*
2013-29 Iwan de Kok (UT), *Listening Heads.*
2013-28 Frans van der Sluis (UT), *When Complexity becomes Interesting: An Inquiry into the Information eXperience.*
2013-27 Mohammad Huq (UT), *Inference-based Framework Managing Data Provenance.*
2013-26 Alireza Zarghami (UT), *Architectural Support for Dynamic Homecare Service Provisioning.*
2013-25 Agnieszka Anna Latoszek-Berendsen (UM), *Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System.*
2013-24 Haitham Bou Ammar (UM), *Automated Transfer in Reinforcement Learning.*
2013-23 Patricio de Alencar Silva (UvT), *Value Activity Monitoring.*
2013-22 Tom Claassen (RUN), *Causal Discovery and Logic.*
2013-21 Sander Wubben (UvT), *Text-to-text generation by monolingual machine translation.*
2013-20 Katja Hofmann (UvA), *Fast and Reliable Online Learning to Rank for Information Retrieval.*
2013-19 Renze Steenhuizen (TUD), *Coordinated Multi-Agent Planning and Scheduling.*
2013-18 Jeroen Janssens (UvT), *Outlier Selection and One-Class Classification.*
2013-17 Koen Kok (VU), *The PowerMatcher: Smart Coordination for the Smart Electricity Grid.*
2013-16 Eric Kok (UU), *Exploring the practical benefits of argumentation in multi-agent deliberation.*
2013-15 Daniel Hennes (UM), *Multiagent Learning - Dynamic Games and Applications.*
2013-14 Jafar Tanha (UVA), *Ensemble Approaches to Semi-Supervised Learning Learning.*
2013-13 Mohammad Safiri (UT), *Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly.*
2013-12 Marian Razavian (VU), *Knowledge-driven Migration to Services.*
2013-11 Evangelos Pournaras (TUD), *Multi-level Reconfigurable Self-organization in Overlay Services.*
2013-10 Jeewanie Jayasinghe Arachchige (UvT), *A Unified Modeling Framework for Service Design.*

2013-09 Fabio Gori (RUN), *Metagenomic Data Analysis: Computational Methods and Applications.*
2013-08 Robbert-Jan Merk(VU), *Making enemies: cognitive modeling for opponent agents in fighter pilot simulators.*
2013-07 Giel van Lankveld (UvT), *Quantifying Individual Player Differences.*
2013-06 Romulo Goncalves (CWI), *The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience.*
2013-05 Dulce Pumareja (UT), *Groupware Requirements Evolutions Patterns.*
2013-04 Chetan Yadati (TUD), *Coordinating autonomous planning and scheduling.*
2013-03 Szymon Klarman (VU), *Reasoning with Contexts in Description Logics.*
2013-02 Erietta Liarou (CWI), *MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing.*
2013-01 Viorel Milea (EUR), *News Analytics for Financial Decision Support.*
2012-51 Jeroen de Jong (TUD), *Heuristics in Dynamic Sceduling; a practical framework with a case study in elevator dispatching.*
2012-50 Steven van Kervel (TUD), *Ontologogy driven Enterprise Information Systems Engineering.*
2012-49 Michael Kaisers (UM), *Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions.*
2012-48 Jorn Bakker (TUE), *Handling Abrupt Changes in Evolving Time-series Data.*
2012-47 Manos Tsagkias (UVA), *Mining Social Media: Tracking Content and Predicting Behavior.*
2012-46 Simon Carter (UVA), *Exploration and Exploitation of Multilingual Data for Statistical Machine Translation.*
2012-45 Benedikt Kratz (UvT), *A Model and Language for Business-aware Transactions.*
2012-44 Anna Tordai (VU), *On Combining Alignment Techniques.*
2012-43 Withdrawn, .
2012-42 Dominique Verpoorten (OU), *Reflection Amplifiers in self-regulated Learning.*
2012-41 Sebastian Kelle (OU), *Game Design Patterns for Learning.*
2012-40 Agus Gunawan (UvT), *Information Access for SMEs in Indonesia.*
2012-39 Hassan Fatemi (UT), *Risk-aware design of value and coordination networks.*
2012-38 Selmar Smit (VU), *Parameter Tuning and Scientific Testing in Evolutionary Algorithms.*
2012-37 Agnes Nakakawa (RUN), *A Collaboration Process for Enterprise Architecture Creation.*
2012-36 Denis Ssebugwawo (RUN), *Analysis and Evaluation of Collaborative Modeling Processes.*
2012-35 Evert Haasdijk (VU), *Never Too Old To Learn - On-line Evolution of Controllers in Swarm- and Modular Robotics.*
2012-34 Pavol Jancura (RUN), *Evolutionary analysis in PPI networks and applications.*
2012-33 Rory Sie (OUN), *Coalitions in Cooperation Networks (COCOON).*
2012-32 Wietske Visser (TUD), *Qualitative multi-criteria preference representation and reasoning.*
2012-31 Emily Bagarukayo (RUN), *A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure.*
2012-30 Alina Pommeranz (TUD), *Designing Human-Centered Systems for Reflective Decision Making.*
2012-29 Almer Tigelaar (UT), *Peer-to-Peer Information Retrieval.*
2012-28 Nancy Pascall (UvT), *Engendering Technology Empowering Women.*
2012-27 Hayrettin Gurkok (UT), *Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games.*
2012-26 Emile de Maat (UVA, *Making Sense of Legal Text.*

2012-25 Silja Eckartz (UT), *Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application.*
2012-24 Laurens van der Werff (UT), *Evaluation of Noisy Transcripts for Spoken Document Retrieval.*
2012-23 Christian Muehl (UT), *Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction.*
2012-22 Thijs Vis (UvT), *Intelligence, politie en veiligheidsdienst: verenigbare grootheden?.*
2012-21 Roberto Cornacchia (TUD), *Querying Sparse Matrices for Information Retrieval.*
2012-20 Ali Bahramisharif (RUN), *Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing.*
2012-19 Helen Schonenberg (TUE), *What's Next? Operational Support for Business Process Execution.*
2012-18 Eltjo Poort (VU), *Improving Solution Architecting Practices.*
2012-17 Amal Elgammal (UvT), *Towards a Comprehensive Framework for Business Process Compliance.*
2012-16 Fiemke Both (VU), *Helping people by understanding them - Ambient Agents supporting task execution and depression treatment.*
2012-15 Natalie van der Wal (VU), *Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes.*
2012-14 Evgeny Knutov (TUE), *Generic Adaptation Framework for Unifying Adaptive Web-based Systems.*
2012-13 Suleman Shahid (UvT), *Fun and Face: Exploring non-verbal expressions of emotion during playful interactions.*
2012-12 Kees van der Sluijs (TUE), *Model Driven Design and Data Integration in Semantic Web Information Systems.*
2012-11 J.C.B. Rantham Prabhakara (TUE), *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics.*
2012-10 David Smits (TUE), *Towards a Generic Distributed Adaptive Hypermedia Environment.*
2012-09 Ricardo Neisse (UT), *Trust and Privacy Management Support for Context-Aware Service Platforms.*
2012-08 Gerben de Vries (UVA), *Kernel Methods for Vessel Trajectories.*
2012-07 Rianne van Lambalgen (VU), *When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions.*
2012-06 Wolfgang Reinhardt (OU), *Awareness Support for Knowledge Workers in Research Networks.*
2012-05 Marijn Plomp (UU), *Maturing Interorganisational Information Systems.*
2012-04 Jurriaan Souer (UU), *Development of Content Management System-based Web Applications.*
2012-03 Adam Vanya (VU), *Supporting Architecture Evolution by Mining Software Repositories.*
2012-02 Muhammad Umair(VU), *Adaptivity, emotion, and Rationality in Human and Ambient Agent Models.*
2012-01 Terry Kakeeto (UvT), *Relationship Marketing for SMEs in Uganda.*
2011-49 Andreea Niculescu (UT), *Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality.*
2011-48 Mark Ter Maat (UT), *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent.*
2011-47 Azizi Bin Ab Aziz(VU), *Exploring Computational Models for Intelligent Support of Persons with Depression.*
2011-46 Beibei Hu (TUD), *Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work.*

2011-45 Herman Stehouwer (UvT), *Statistical Language Models for Alternative Sequence Selection.*
2011-44 Boris Reuderink (UT), *Robust Brain-Computer Interfaces.*
2011-43 Henk van der Schuur (UU), *Process Improvement through Software Operation Knowledge.*
2011-42 Michal Sindlar (UU), *Explaining Behavior through Mental State Attribution.*
2011-41 Luan Ibraimi (UT), *Cryptographically Enforced Distributed Data Access Control.*
2011-40 Viktor Clerc (VU), *Architectural Knowledge Management in Global Software Development.*
2011-39 Joost Westra (UU), *Organizing Adaptation using Agents in Serious Games.*
2011-38 Nyree Lemmens (UM), *Bee-inspired Distributed Optimization.*
2011-37 Adriana Burlutiu (RUN), *Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference.*
2011-36 Erik van der Spek (UU), *Experiments in serious game design: a cognitive approach.*
2011-35 Maaike Harbers (UU), *Explaining Agent Behavior in Virtual Training.*
2011-34 Paolo Turrini (UU), *Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations.*
2011-33 Tom van der Weide (UU), *Arguing to Motivate Decisions.*
2011-32 Nees-Jan van Eck (EUR), *Methodological Advances in Bibliometric Mapping of Science.*
2011-31 Ludo Waltman (EUR), *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality.*
2011-30 Egon van den Broek (UT), *Affective Signal Processing (ASP): Unraveling the mystery of emotions.*
2011-29 Faisal Kamiran (TUE), *Discrimination-aware Classification.*
2011-28 Rianne Kaptein (UVA), *Effective Focused Retrieval by Exploiting Query Context and Document Structure.*
2011-27 Aniel Bhulai (VU), *Dynamic website optimization through autonomous management of design patterns.*
2011-26 Matthijs Aart Pontier (VU), *Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots.*
2011-25 Syed Waqar ul Qounain Jaffry (VU)), *Analysis and Validation of Models for Trust Dynamics.*
2011-24 Herwin van Welbergen (UT), *Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior.*
2011-23 Wouter Weerkamp (UVA), *Finding People and their Utterances in Social Media.*
2011-22 Junte Zhang (UVA), *System Evaluation of Archival Description and Access.*
2011-21 Linda Terlouw (TUD), *Modularization and Specification of Service-Oriented Systems.*
2011-20 Qing Gu (VU), *Guiding service-oriented software engineering - A view-based approach.*
2011-19 Ellen Rusman (OU), *The Mind's Eye on Personal Profiles.*
2011-18 Mark Ponsen (UM), *Strategic Decision-Making in complex games.*
2011-17 Jiyin He (UVA), *Exploring Topic Structure: Coherence, Diversity and Relatedness.*
2011-16 Maarten Schadd (UM), *Selective Search in Games of Different Complexity.*
2011-15 Marijn Koolen (UvA), *The Meaning of Structure: the Value of Link Evidence for Information Retrieval.*
2011-14 Milan Lovric (EUR), *Behavioral Finance and Agent-Based Artificial Markets.*
2011-13 Xiaoyu Mao (UvT), *Airport under Control. Multiagent Scheduling for Airport Ground Handling.*
2011-12 Carmen Bratosin (TUE), *Grid Architecture for Distributed Process Mining.*
2011-11 Dhaval Vyas (UT), *Designing for Awareness: An Experience-focused HCI Perspective.*

2011-10 Bart Bogaert (UvT), *Cloud Content Contention.*
2011-09 Tim de Jong (OU), *Contextualised Mobile Media for Learning.*
2011-08 Nieske Vergunst (UU), *BDI-based Generation of Robust Task-Oriented Dialogues.*
2011-07 Yujia Cao (UT), *Multimodal Information Presentation for High Load Human Computer Interaction.*
2011-06 Yiwen Wang (TUE), *Semantically-Enhanced Recommendations in Cultural Heritage.*
2011-05 Base van der Raadt (VU), *Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.*
2011-04 Hado van Hasselt (UU), *Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference.*
2011-03 Jan Martijn van der Werf (TUE), *Compositional Design and Verification of Component-Based Information Systems.*
2011-02 Nick Tinnemeier (UU), *Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language.*
2011-01 Botond Cseke (RUN), *Variational Algorithms for Bayesian Inference in Latent Gaussian Models.*
2010-53 Edgar Meij (UVA), *Combining Concepts and Language Models for Information Access.*
2010-52 Peter-Paul van Maanen (VU), *Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention.*
2010-51 Alia Khairia Amin (CWI), *Understanding and supporting information seeking tasks in multiple sources.*
2010-50 Bouke Huurnink (UVA), *Search in Audiovisual Broadcast Archives.*
2010-49 Jahn-Takeshi Saito (UM), *Solving difficult game positions.*
2010-48 Withdrawn, .
2010-47 Chen Li (UT), *Mining Process Model Variants: Challenges, Techniques, Examples.*
2010-46 Vincent Pijpers (VU), *e3alignment: Exploring Inter-Organizational Business-ICT Alignment.*
2010-45 Vasilios Andrikopoulos (UvT), *A theory and model for the evolution of software services.*
2010-44 Pieter Bellekens (TUE), *An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain.*
2010-43 Peter van Kranenburg (UU), *A Computational Approach to Content-Based Retrieval of Folk Song Melodies.*
2010-42 Sybren de Kinderen (VU), *Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach.*
2010-41 Guillaume Chaslot (UM), *Monte-Carlo Tree Search.*
2010-40 Mark van Assem (VU), *Converting and Integrating Vocabularies for the Semantic Web.*
2010-39 Ghazanfar Farooq Siddiqui (VU), *Integrative modeling of emotions in virtual agents.*
2010-38 Dirk Fahland (TUE), *From Scenarios to components.*
2010-37 Niels Lohmann (TUE), *Correctness of services and their composition.*
2010-36 Jose Janssen (OU), *Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification.*
2010-35 Dolf Trieschnigg (UT), *Proof of Concept: Concept-based Biomedical Information Retrieval.*
2010-34 Teduh Dirgahayu (UT), *Interaction Design in Service Compositions.*
2010-33 Robin Aly (UT), *Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval.*
2010-32 Marcel Hiel (UvT), *An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems.*
2010-31 Victor de Boer (UVA), *Ontology Enrichment from Heterogeneous Sources on the Web.*
2010-30 Marieke van Erp (UvT), *Accessing Natural History - Discoveries in data cleaning, struc-*

turing, and retrieval.
2010-29 Stratos Idreos (CWI), *Database Cracking: Towards Auto-tuning Database Kernels.*
2010-28 Arne Koopman (UU), *Characteristic Relational Patterns.*
2010-27 Marten Voulon (UL), *Automatisch contracteren.*
2010-26 Ying Zhang (CWI), *XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines.*
2010-25 Zulfiqar Ali Memon (VU), *Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective.*
2010-24 Dmytro Tykhonov, *Designing Generic and Efficient Negotiation Strategies.*
2010-23 Bas Steunebrink (UU), *The Logical Structure of Emotions.*
2010-22 Michiel Hildebrand (CWI), *End-user Support for Access to Heterogeneous Linked Data.*
2010-21 Harold van Heerde (UT), *Privacy-aware data management by means of data degradation.*
2010-20 Ivo Swartjes (UT), *Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative.*
2010-19 Henriette Cramer (UvA), *People's Responses to Autonomous and Adaptive Systems.*
2010-18 Charlotte Gerritsen (VU), *Caught in the Act: Investigating Crime by Agent-Based Simulation.*
2010-17 Spyros Kotoulas (VU), *Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications.*
2010-16 Sicco Verwer (TUD), *Efficient Identification of Timed Automata, theory and practice.*
2010-15 Lianne Bodenstaff (UT), *Managing Dependency Relations in Inter-Organizational Models.*
2010-14 Sander van Splunter (VU), *Automated Web Service Reconfiguration.*
2010-13 Gianluigi Folino (RUN), *High Performance Data Mining using Bio-inspired techniques.*
2010-12 Susan van den Braak (UU), *Sensemaking software for crime analysis.*
2010-11 Adriaan Ter Mors (TUD), *The world according to MARP: Multi-Agent Route Planning.*
2010-10 Rebecca Ong (UL), *Mobile Communication and Protection of Children.*
2010-09 Hugo Kielman (UL), *A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging.*
2010-08 Krzysztof Siewicz (UL), *Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments.*
2010-07 Wim Fikkert (UT), *Gesture interaction at a Distance.*
2010-06 Sander Bakkes (UvT), *Rapid Adaptation of Video Game AI.*
2010-05 Claudia Hauff (UT), *Predicting the Effectiveness of Queries and Retrieval Systems.*
2010-04 Olga Kulyk (UT), *Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments.*
2010-03 Joost Geurts (CWI), *A Document Engineering Model and Processing Framework for Multimedia documents.*
2010-02 Ingo Wassink (UT), *Work flows in Life Science.*
2010-01 Matthijs van Leeuwen (UU), *Patterns that Matter.*
2009-46 Loredana Afanasiev (UvA), *Querying XML: Benchmarks and Recursion.*
2009-45 Jilles Vreeken (UU), *Making Pattern Mining Useful.*
2009-44 Roberto Santana Tapia (UT), *Assessing Business-IT Alignment in Networked Organizations.*
2009-43 Virginia Nunes Leal Franqueira (UT), *Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients.*
2009-42 Toine Bogers (UvT), *Recommender Systems for Social Bookmarking.*
2009-41 Igor Berezhnyy (UvT), *Digital Analysis of Paintings.*
2009-40 Stephan Raaijmakers (UvT), *Multinomial Language Learning: Investigations into the*

Geometry of Language.
2009-39 Christian Stahl (TUE, Humboldt-Universitaet zu Berlin), *Service Substitution – A Behavioral Approach Based on Petri Nets.*
2009-38 Riina Vuorikari (OU), *Tags and self-organisation: a metadata ecology for learning resources in a multilingual context.*
2009-37 Hendrik Drachsler (OUN), *Navigation Support for Learners in Informal Learning Networks.*
2009-36 Marco Kalz (OUN), *Placement Support for Learners in Learning Networks.*
2009-35 Wouter Koelewijn (UL), *Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling.*
2009-34 Inge van de Weerd (UU), *Advancing in Software Product Management: An Incremental Method Engineering Approach.*
2009-33 Khiet Truong (UT), *How Does Real Affect Affect Affect Recognition In Speech?.*
2009-32 Rik Farenhorst (VU) and Remco de Boer (VU), *Architectural Knowledge Management: Supporting Architects and Auditors.*
2009-31 Sofiya Katrenko (UVA), *A Closer Look at Learning Relations from Text.*
2009-30 Marcin Zukowski (CWI), *Balancing vectorized query execution with bandwidth-optimized storage.*
2009-29 Stanislav Pokraev (UT), *Model-Driven Semantic Integration of Service-Oriented Applications.*
2009-28 Sander Evers (UT), *Sensor Data Management with Probabilistic Models.*
2009-27 Christian Glahn (OU), *Contextual Support of social Engagement and Reflection on the Web.*
2009-26 Fernando Koch (UU), *An Agent-Based Model for the Development of Intelligent Mobile Services.*
2009-25 Alex van Ballegooij (CWI), *"RAM: Array Database Management through Relational Mapping".*
2009-24 Annerieke Heuvelink (VUA), *Cognitive Models for Training Simulations.*
2009-23 Peter Hofgesang (VU), *Modelling Web Usage in a Changing Environment.*
2009-22 Pavel Serdyukov (UT), *Search For Expertise: Going beyond direct evidence.*
2009-21 Stijn Vanderlooy (UM), *Ranking and Reliable Classification.*
2009-20 Bob van der Vecht (UU), *Adjustable Autonomy: Controling Influences on Decision Making.*
2009-19 Valentin Robu (CWI), *Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets.*
2009-18 Fabian Groffen (CWI), *Armada, An Evolving Database System.*
2009-17 Laurens van der Maaten (UvT), *Feature Extraction from Visual Data.*
2009-16 Fritz Reul (UvT), *New Architectures in Computer Chess.*
2009-15 Rinke Hoekstra (UVA), *Ontology Representation - Design Patterns and Ontologies that Make Sense.*
2009-14 Maksym Korotkiy (VU), *From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA).*
2009-13 Steven de Jong (UM), *Fairness in Multi-Agent Systems.*
2009-12 Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin), *Operating Guidelines for Services.*
2009-11 Alexander Boer (UVA), *Legal Theory, Sources of Law & the Semantic Web.*
2009-10 Jan Wielemaker (UVA), *Logic programming for knowledge-intensive interactive applications.*
2009-09 Benjamin Kanagwa (RUN), *Design, Discovery and Construction of Service-oriented Systems.*
2009-08 Volker Nannen (VU), *Evolutionary Agent-Based Policy Analysis in Dynamic Environ-*

ments.
2009-07 Ronald Poppe (UT), *Discriminative Vision-Based Recovery and Recognition of Human Motion.*
2009-06 Muhammad Subianto (UU), *Understanding Classification.*
2009-05 Sietse Overbeek (RUN), *Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality.*
2009-04 Josephine Nabukenya (RUN), *Improving the Quality of Organisational Policy Making using Collaboration Engineering.*
2009-03 Hans Stol (UvT), *A Framework for Evidence-based Policy Making Using IT.*
2009-02 Willem Robert van Hage (VU), *Evaluating Ontology-Alignment Techniques.*
2009-01 Rasa Jurgelenaite (RUN), *Symmetric Causal Independence Models.*