A Computational Model of Non-Cooperation in Natural Language Dialogue

Thesis

A Computational Model of Non-Cooperation in Natural Language Dialogue

Brian Plüss
Bachiller Universitario en Ciencias Exactas
Licenciado en Ciencias de la Computación

Computing and Communications Department
The Open University

A thesis submitted for the degree of
Doctor of Philosophy

September 2013

Abstract

A common assumption in the study of conversation is that participants fully cooperate in order to maximise the effectiveness of the exchange and ensure communication flow. This assumption persists even in situations in which the private goals of the participants are at odds: they may act strategically pursuing their agendas, but will still adhere to a number of linguistic norms or conventions which are implicitly accepted by a community of language users.

However, in naturally occurring dialogue participants often depart from such norms, for instance, by asking inappropriate questions, by avoiding to provide adequate answers or by volunteering information that is not relevant to the conversation. These are examples of what we call linguistic non-cooperation.

This thesis presents a systematic investigation of linguistic non-cooperation in dialogue. Given a specific activity, in a specific cultural context and time, the method proceeds by making explicit which linguistic behaviours are appropriate. This results in a set of rules: the global dialogue game. Non-cooperation is then measured as instances in which the actions of the participants are not in accordance with these rules. The dialogue game is formally defined in terms of discourse obligations. These are actions that participants are expected to perform at a given point in the dialogue based on the dialogue history. In this context, non-cooperation amounts to participants failing to act according to their obligations.

We propose a general definition of linguistic non-cooperation and give a specific instance for political interview dialogues. Based on the latter, we

present an empirical method which involves a coding scheme for the manual annotation of interview transcripts. The degree to which each participant cooperates is automatically determined by contrasting the annotated transcripts with the rules in the dialogue game for political interviews. The approach is evaluated on a corpus of broadcast political interviews and tested for correlation with human judgement on the same corpus.

Further, we describe a model of conversational agents that incorporates the concepts and mechanisms above as part of their dialogue manager. This allows for the generation of conversations in which the agents exhibit varying degrees of cooperation by controlling how often they favour their private goals instead of discharging their discourse obligations.

Acknowledgements

These pages close a very long journey, the last part of which started in 2008 when I left Argentina and moved to Milton Keynes in the UK. I embarked on a PhD as a natural step in my education, but also as a personal challenge. And what a challenge it was! The experience was thrilling at times, but also extremely difficult. It would have been impossible to reach this point without an army of people that played essential roles in different and unique ways. To them I am deeply grateful and the following lines are my attempt to put this in writing. Still, words are nowhere nearly enough to communicate how important they have been over the years.

First and foremost, I thank my supervisors, Paul Piwek and Richard Power, for their constant guidance, encouragement and understanding. They received a computer scientist with a formal background and a huge lack of knowledge in computational linguistics and took on the challenge to turn him into an independent researcher in an area that builds upon empirical evidence. Along the way, they coped with my Latin American manners and patiently taught me to listen to questions and suggestions before attempting a reply. They gave me enough freedom to do exactly the research project I wanted, but knew when to show me where I was going down potentially unfruitful paths. They taught me how to write, leaving "woolly generalisations" for café talks. I have immense respect for them and gratitude for the time they have invested in me. I hope to do them justice in the next stages of my academic and professional life.

The Open University is a great environment to study and work. Its unique mission of spreading higher education to "all who wish to realise their ambitions and fulfil their potential" is beautiful and hugely necessary. It fills me with pride to be part of this institution. I am thankful to the Faculty of Mathematics, Computing and Technology for funding my studies, especially into the initially unplanned fourth year. Many academics in the Department of Computing and Communications played an essential role in my progress. Marian Petre and Robin Laney dedicated much time and effort in teaching me and my colleagues how to do good research. I thank Marian for having taken an exaggerate amount of time to push me forward when the task ahead looked impossible. Also Michel Wermelinger, Leonor Barroca, Anne De Roeck and Alistair Willis helped at different stages along the way. I am thankful to Hugh Robinson who "grilled" me the first time I presented my research, with questions that ultimately shaped the project to a great extent. The members, past and present, of the Natural Language Generation group were "home", academically speaking, and provided a safe environment where I could present my often unpolished ideas. They are Donia Scott, Sandra Williams, Svetlana Stoyanchev, Allan Third, Eva Banik, Richard Doust, Sharon Moyo and Tu Anh Nguyen.

Half-way through my studies, I was immensely lucky to take part in a summer internship at the University of Southern California's Institute for Creative Technologies. Working under the supervision of David DeVault, with David Traum and the rest of the Natural Language Dialogue Group, was an excellent opportunity, professionally and also personally. I also thank David DeVault, David Traum, Ron Artstein and Kalliroi Georgila for illuminating discussions of my own research.

Milton Keynes would not have been the happy memory it is today without my dear friends, who would turn any city into the best place to be. I thank Bethany for her constant company during the first three years while we shared 36 Warren Bank. She turned an otherwise uneventful flat into a fairground full of excitement and endless fun, as we built a unique connection that will last for the rest of our lives. I thank Koula and Jo for being the other two legs of that sturdy table of friendship and mutual understanding. It is impossible for me to name everyone else in the Milton Keynes gang without forgetting someone. I implicitly and wholeheartedly thank all of them and here attempt a necessarily incomplete list: Sam, Dave, Lionel, Andrés, Katarina, Davey, Paul W, Ben R, Ben D, Marc, Andy C, Andy M, Adam, Miriam, Jakub, Dinar, Giorgos, Mariano, Lucas, Stephen, Minh, Tu Anh, Laura, Stefan, Macarena, Joan, Lana, Alexandra, Nadia, Jen, Anna, Carlos, Sally, Eva, Alba, Thomas, Saad, Vanessa, Philippe, Liliana, Carmen, Rob, Beck, Amanda, Zdenek, Mic, Carolyn... And Paul T.

I also thank those friends elsewhere in the UK: Alejandra, Leandro, Ros, Duncan, Charlie, Andy, Sergio, Santiago, Tasha and Kirsty. And those around the world: Ana, Jõao, Tiago, Cristina, Francisca, Fernando, Pablo and Luis in Portugal; Quitterie, Jano and Nache in France; Abdul, Matias, Virpa, Peikko and Em in Finland; Aly and Rachel in the US; Damián and Mercedes in Italy; and about a thousand in Argentina. I am thankful to Maxi, my teacher and friend, for opening so many doors and for recent collaborations which will hopefully continue for years to come. I thank those who opened their homes and provided a safe haven for me to work during the long writing up stage: Bethany and Dave, Miriam and Jakub, Alejandra and Leandro, Laura and Luciano, Chicho and Silvina. Diego not only opened his home, but also allowed for a profound connection that has been essential all along, especially in turbulent times. Over the years and despite the ocean in between, Dani remained that soulmate I can go to, whatever the matter and without fear of judgement or incomprehension.

Dermot showed up right before things started to get complicated and walked with me through the hardest last two years. Without his company, understanding and support – in every imaginable way – I would have certainly failed. A lifetime will not be enough to pay back his immense effort.

And finally, I need to thank my beloved family. My parents, Gabriel and Graziela, endured the distance and were there whenever I needed them, without necessarily understanding what I was working on and why I was a student well into my thirties. I thank my brother, Alan, the architect and traveller who in these years grew up into a man with an amazingly wise and soothing approach to life. My grandmother, Raquel, endless in her wisdom, has been my life coach for as many years as I can remember. Our long and frequent conversations have shown me the way out of many dark alleys, and they still do today. My aunts and uncles and my many cousins were a colourful and beautifully noisy group to go back to when visiting Argentina. I am thankful to María José, my cousin-sister, so full of much-needed happiness and fresh air.

There are four people without whom reaching the end of this project would have been impossible: my parents, my grandmother and Dermot. This thesis is dedicated to them, in infinite gratitude.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Most approaches to modelling conversation are based on a notion of full cooperation between the dialogue participants. Traditional models relying on intentions (Grosz and Sidner, 1986; Cohen and Levesque, 1991), conversational games (Power, 1979; Carletta et al., 1997), shared plans (Grosz and Sidner, 1990; Chu-Carroll and Carberry, 1998) or collaborative problem-solving (Blaylock and Allen, 2005) explain dialogue situations in which participants recognise each other's intentions and, at least to some extent, adopt each other's goals when deciding on their actions. These assumptions are theoretically grounded, as most work in linguistics has considered situations in which participants share a common goal and cooperate to achieve it by means of conversation (Grice, 1975; Clark and Schaefer, 1989). They are also practically sound: dialogue models are usually implemented in the form of dialogue systems, built for the purpose of providing a service to their users. In this scenario, failure to cooperate either on the side of the system or of the user is against the premises on which the system is conceived and used.

In everyday conversation, however, a great many situations do not conform to these assumptions. Consider the example below, in which BBC presenter Jeremy Paxman questions former UK Home Secretary Michael Howard with respect to a meeting in 1995 between Howard and the head of the Prison Service, Derek Lewis, about the dismissal of the governor of Parkhurst Prison, John Marriott, due to repeated security failures. The case was given considerable attention in the media, as a result of accusations by Lewis that Howard had instructed him, thus exceeding the powers of his office[1]:

Example 1.1.

PAXMAN:	(...) Are you saying Mr Lewis is lying?
HOWARD:	I have given a full account of this, and the position is what I told the House of Commons, and let me tell you what the position is-
PAXMAN:	(interrupting) So you are saying that Mr Lewis lied?
HOWARD:	Let me tell you exactly what the position is. I was entitled to be consulted and I was consulted, I was entitled to express an opinion and I did express an opinion. I was not entitled to instruct Derek Lewis what to do, and I did not instruct him what to do and you will understand and recall that Mr Marriot was not suspended, he was moved, and Derek Lewis told the Select Committee of the House of Commons that it was his opinion, Derek Lewis's opinion, that he should be moved immediately. That is what happened.
PAXMAN:	Mr Lewis says: I – that is Mr Lewis – told him what we had decided about Marriot, and why; he – that is you – exploded; simply moving the governor was politically unpalatable, it sounded indecisive, it would be seen as a fudge; if I did not change my mind and suspend Marriot he would have to consider overruling me. You can't both be right.
HOWARD:	Mr Marriot was not suspended. I was entitled to express my views. I was entitled to be consulted-
PAXMAN:	(interrupting) Did you threaten to overrule him?

[1] An extended fragment of this interview can be found in http://www.youtube.com/watch?v=Uwlsd8RAoqI (last accessed: September 2013).

HOWARD:	I was not entitled to instruct Derek Lewis, and I did not instruct him.
PAXMAN:	Did you threaten to overrule him?
HOWARD:	The truth of the matter is that Mr. Marriott was not suspended. I-
PAXMAN:	(overlapping) Did you threaten to overrule him?
HOWARD:	-did not overrule Derek Lewis.
PAXMAN:	Did you threaten to overrule him?
HOWARD:	I took advice on what I could or could not do-
PAXMAN:	(overlapping) Did you threaten to overrule him, Mr. Howard?
HOWARD:	-and I acted scrupulously in accordance with that advice, I did not overrule Derek Lewis-
PAXMAN:	(overlapping) Did you threaten to overrule him?
HOWARD:	- Mr. Marriott was not suspended.
PAXMAN:	Did you threaten to overrule him?
HOWARD:	I have accounted for my decision to dismiss Derek Lewis-
PAXMAN:	(overlapping) Did you threaten to overrule him?
HOWARD:	-in great detail, before the House of Commons.
PAXMAN:	I note that you're not answering the question of whether you threatened to overrule him.

(*Newsnight*, BBC, 1997)

While at some level Paxman and Howard are sharing a goal, for otherwise they would not be having an interview, the exchange is clearly conflictive, to the point that their behaviour compromises the flow of the conversation.

The fragment below took place seven years after the exchange in Example 1.1, when public awareness of the 1995 affair had dissipated:

Example 1.2.

PAXMAN:	Can you clear up whether or not you did threaten to overrule Derek Lewis when you were Home Secretary?
HOWARD:	Oh, come on, Jeremy, you are really going to go back over that again? As...
PAXMAN:	(overlapping) You've had seven years to think about it!
HOWARD:	(overlapping)... as, as it happens, I didn't. Are you satisfied now?
PAXMAN:	Thank you. Why didn't you say that at the time?

HOWARD: I, well, we've been over this many, many times. I, I, I knew that everyone was crawling over every syllable I said about that, and I wanted to check very carefully what I said before answering your question.

(*Newsnight*, BBC, 2004)

On this occasion, Howard provides an answer almost immediately and the flow of the conversation contrasts noticeably with that in Example 1.1.

Below is another example. Jeremy Paxman interviews British MP George Galloway, shortly after his victory in the UK 2005 General Election[2]:

Example 1.3.

PAXMAN: We're joined now from his count in Bethnal Green and Bow by George Galloway. Mr Galloway, are you proud of having got rid of one of the very few black women in Parliament?

GALLOWAY: What a preposterous question. I know it's very late in the night, but wouldn't you be better starting by congratulating me for one of the most sensational election results in modern history?

PAXMAN: Are you proud of having got rid of one of the very few black women in Parliament?

GALLOWAY: I'm not- Jeremy, move on to your next question.

PAXMAN: You're not answering that one?

GALLOWAY: No, because I don't believe that people get elected because of the colour of their skin. I believe people get elected because of their record and because of their policies. So move on to your next question.

PAXMAN: Are you proud-

GALLOWAY: (Interrupting) Because I've got a lot of people who want to speak to me. If you ask that question again, I'm going, I warn you now.

PAXMAN: Don't try and threaten me Mr Galloway, please. (...)

GALLOWAY: You are actually conducting one of the most - even by your standards - one of the most absurd interviews I have ever participated in. I have just won an election. Can you find it within yourself to recognise that fact? To recognise the fact that the people of Bethnal Green and Bow chose me this evening. Why are you insulting them?

[2]The interview was aired live on 6 May, 2005 and can be found at `http://www.youtube.com/watch?v=dKDuhGOqr8E` (last accessed: September 2013).

PAXMAN:	I'm not insulting them, I'm not insulting you
GALLOWAY:	You are insulting them, they chose me just a few minutes ago. Can't you find it within yourself even to congratulate me on this victory?
PAXMAN:	Congratulations, Mr Galloway.
GALLOWAY:	Thank you very much indeed. *(Waves, removes microphone and leaves)*

(*UK General Election*, BBC, 2005)

This exchange too differs noticeably from typical political interviews, in which one of the participants poses more or less impartial questions, while the other provides clear and relevant answers. The interaction is confrontational from the outset, to the point that the interviewee eventually abandons the conversation.

The investigation reported in this thesis aims at shedding light on the nature of non-cooperation in dialogue, by capturing the intuitions that allow us to distinguish the conversational behaviour of the participants in interactions like in Examples 1.1 and 1.3 from those like in Example 1.2, with respect to how a dialogue of a certain type – in this case a political interview – should normally go. Heritage describes the distinctive roles of participants in news interviews as follows (Heritage, 1998, p. 8):

> "the participants -IRs [=interviewers] and IEs [=interviewees]- exclude themselves from a wide variety of actions that they are normally free to do in the give and take of ordinary conversation. If IRs restrict themselves to asking questions, then they cannot – at least overtly – express opinions, or argue with, debate or criticize the interviewees' positions nor, conversely, agree with, support or defend them. Correspondingly, if IEs restrict themselves to answers (or responses) to questions, then they cannot ask questions (of IRs or other IEs), nor make unsolicited com-

> ments on previous remarks, initiate changes of topic, or divert the discussion into criticisms of the IR or the broadcasting organization."

Walton and Krabbe (1995) in their study of argumentation dialogues formalise dialogue types by means of precise rules. They conclude that rigorous models of conversational interaction are useful analytical tools, but accept that most of the huge variety of everyday conversation escapes it. Such characterisations are often based on strict rules that capture typical dialogue situations while leaving out considerable detail. As the examples above show, actual participant behaviour can diverge from the typical case in unexpected ways, falling outside such characterisations[3]. It could be argued that it is always possible to account for interactions like those in Examples 1.1 and 1.3 by adding further rules to capture the variations present in these conversations. Still, in the limit this approach would require an additional set of rules for each possible unconventional behaviour.

At the same time, the rules and patterns captured by formal models are useful also in exceptional cases. As these models describe expected or permissible behaviour in a certain conversational scenario, they provide a basis against which actual behaviour can be assessed in order to detect deviations. This research aims at reconciling two worlds, using the insights from formal models as descriptions of expected, conventional behaviour in the form of social obligations, but looking at naturally-occurring cases that deviate from the norm. This, in turn, calls for the definition of non-cooperative conversational behaviour and for the techniques to detect this accurately and reliably, which are at the core of our contribution.

[3]Consider, for instance, Ginzburg's QUD model (Ginzburg, 1996) when applied to dialogue (1), in which Howard repeatedly fails to either accept or reject Paxman's question.

1.1 What is Non-Cooperation in Dialogue?

During a conversation, participants interact in many ways. They speak and listen, decode each other's utterances, signal understanding, or request clarifications. They request and provide information, accept questions and statements, or point out inadequate proposals and offer alternatives. These are all linguistic tasks that keep the conversation flowing.

In most cases, conversation supports a social activity that determines how these actions ought to be performed. Successful conversation contributes to the aims of the social activity. At this level a second layer of interaction is observed, in which participants ask useful questions, provide truthful information and generally work together towards the completion of the activity. The latter requires that the individual goals that the participants bring to the conversation align without conflict. When this is not the case, non-cooperation arises. Participants can fail to cooperate at the level of the social activity, for instance by not providing the information that the other party needs. This may or may not translate to non-cooperation at the linguistic level, for instance by remaining totally silent or by diverting the course of the conversation without stating the reasons for rejecting the question.

However, participants can still cooperate at the conversational level without contributing to the goals of the social activity. Consider, for instance, a witness under interrogation in a U.S. trial refusing to answer a question by appealing to the Fifth Amendment of the Constitution[4]. Such behaviour will be accepted in the conversational setting as established by law, although it is not cooperative in relation with the goals of the trial. A

[4] "*No person shall (...) be compelled in any criminal case to be a witness against himself*".

linguistically non-cooperative alternative would be the same witness remaining silent, rather than answering or appealing to the Fifth Amendment. To illustrate further, consider a fictional alternative to the exchange in Example 1.1, where Howard replies by saying "I will not answer that question, as it is not relevant to whether I exceeded the powers of my office". This is still not cooperative for the goals of the interview as it is not contributing information that could be interesting to the audience, but it is cooperative at the linguistic level. It would help in preserving the flow of the conversation, e.g., by triggering a sub-dialogue to solve the disagreement.

The above calls for a treatment of cooperation at two clearly distinctive levels of interaction: that of linguistic actions and that of social or task-related activities. This distinction has been addressed before. Attardo (1997) revisits Gricean pragmatics, relating non-linguistic cooperation to participants' behaviour towards realising task-related goals, and linguistic cooperation to assumptions on their respective behaviour in order to encode and decode intended meaning. From a computational perspective, Bunt (1994) relies on a similar distinction for defining dialogue acts. Also, Traum and Allen (1994) introduce discourse obligations as an alternative to joint intentions and shared plans, to allow for models of dialogues in which participants do not share the same high-level goals and where behaviour is also determined by "a sense of obligation to behave within limits set by the society" (Traum and Allen, 1994, p. 2). Walton and Krabbe (1995) propose a typology of dialogue based on the initial situation triggering the exchange and participants' shared aims and individual goals. Based on their work, Reed and Long (1997) distinguish cases where participants follow a common set of dialogue rules and stay within a mutually acknowledged framework from a stronger notion in which their individual goals are in the same direction.

The research in this thesis is about linguistic cooperation, understood as participants following the discourse obligations imposed upon them by the social activity in which they are engaged. We do not deal with cooperation at the task-level directly and the main focus is on cases in which dialogue participants purposefully fail to follow their obligations, regardless of their motivations for doing so.

1.2 An Extended Example

The approach we will develop in the central chapters of the thesis roughly amounts to identifying a set of features that distinguish cooperative from non-cooperative linguistic actions. The extent to which a dialogue participant is non-cooperative will then be related to the number of such features with respect to the size of the participant's contributions.

As an approximation of the technique we propose in full later, consider the examples of linguistic misbehaviour listed in Table 1.1, grouped following three aspects of conversation: turn-taking, joint projects and speech acts. We call these non-cooperative features (NCFs). The number of occurrences of these features will determine the degree of non-cooperation (DNC) of an exchange.

Turn-taking rules (Sacks et al., 1974) establish that speakers make their contributions at adequate places and in particular ways. Interlocutors in a political interview are expected to respect transition-relevance places, openings and closings according to social conventions.

Joint projects (Clark, 1996) refer to participants' accepting or rejecting each other's proposals. In political interviews a question can be accepted explicitly or implicitly by providing a direct answer, and rejected explicitly

Turn-Taking	For both speakers: • interrupting • overlapping • ending the exchange abruptly
Joint Projects	Interviewer fails to either: • accept answer • ask next relevant question • move to next topical issue • state irrelevance of answer Interviewee fails to either: • accept question • give relevant answer • reject question
Speech Acts	Interviewer either: • expresses personal opinion • argues, debates with or criticises interviewee's position subjectively • agrees with, supports or defends interviewee's position subjectively Interviewee either: • asks (non-CR) question • makes irrelevant comment • initiates change of topic • criticises interviewer

Table 1.1: Some non-cooperative features for political interviews

by stating how it fails to focus on matters of relevance. Likewise, replies can be accepted explicitly or implicitly by asking a next relevant question or by moving on to a new topical issue.

Speech Act theory (Searle, 1979) classifies utterances according to their associated force and propositional content. Going back to Heritage's comment, in a political interview participants can fail to restrict their speech acts to the force and content expected for their role. Non-cooperative features related to speech acts include the interviewer expressing a personal opinion or criticising subjectively the interviewee's positions, and the interviewee asking questions (except for clarification requests) or making irrelevant comments.

As an example, consider another fragment of the Paxman-Howard interview introduced in Example 1.1, annotated with NCFs (**O**: overlap; **JPF**: joint project failure; **UC**: unsolicited comment; **I**: interruption; **TC**: topic change):

(1)	P[11] :	Uir.1	(overlapping) Did you threaten to overrule him?	**O**
	H[12] :	Uie.1	. . . Mr. Marriot was *not* suspended.	**JPF**
	P[13] :	Uir.2	Did you threaten to overrule him?	**JPF**
	H[14] :	Uie.2	(pauses) I have accounted for my decision to dismiss Derek Lewis. . .	
	P[15] :	Uir.3	(overlapping) Did you threaten to overrule him?	**O**
	H[16] :	Uie.2	. . . in great detail before the House of Commons.	**UC**
	P[17] :	Uir.4	I note that you're not answering the question whether you *threatened* to overrule him.	
	H[18] :	Uie.3	Well, the important aspect of this which it's very clear to bear in mind. . .	**JPF**
	P[19] :	Uir.5	(interrupting) I'm sorry, I'm going to be frightfully rude but. . .	**I**
	H[20] :	Uie.4	Yes, you can. . .	
	P[21] :	Uir.6	(overlapping) I'm sorry. . .	**O**
	H[22] :	Uie.4	(overlapping) . . . you can put the question and I will give you, I will give you an answer.	**O**
	P[23] :	Uir.7	. . . it's a straight yes-or-no question and a straight yes-or-no answer:	
		Uir.8	did you threaten to overrule him?	
	H[24] :	Uie.5	I discussed the matter with Derek Lewis.	
		Uie.6	I gave him the benefit of my opinion.	
		Uie.7	I gave him the benefit of my opinion in strong language, but I did not instruct him because I was not, er, entitled to instruct him.	**UC**
		Uie.8	I was entitled to express my opinion and that is what I did.	**UC**
	P[25] :	Uir.9	With respect, that is not answering the question of whether you threatened to overrule him.	
	H[26] :	Uie.9	It's dealing with the relevant point which was what I was entitled to do and what I was not entitled to do,	**TC**
		Uie.10	and I have dealt with this in detail before the House of Commons and before the select committee.	**UC**

For each participant, the degree of non-cooperation (DNC) is computed as the proportion of utterances with one of more occurrences of these non-cooperative features. Table 1.2 summarises non-cooperative features, utterances and the degree of non-cooperation for each participant.

The core of this thesis is dedicated to providing precise definitions for the

	Paxman (ir)	Howard (ie)
Interruptions	1	0
Overlaps	3	1
Project Failures	1	2
Unsolicited Comments	0	4
Topic Change	0	1
Total NCFs	5	8
Utterances	9	10
DNC	**0.56**	**0.80**

Table 1.2: Non-cooperation in the Paxman-Howard interview

concepts above, to proposing and evaluating sound mechanisms for reliably measuring linguistic non-cooperation in dialogue and to devising a model of conversational agents that can incorporate these elements in conversation.

1.3 Research Question

Our research question reads as follows:

RQ: What elements are needed in a computational model of conversational agents so that they can exhibit and cope with non-cooperative as well as cooperative linguistic behaviour in dialogue, in particular in the domain of political interviews?

This research belongs in the area of computational pragmatics: the study of language in use or language in context (Levinson, 1983) from a computational perspective. This involves finding computational models of phenomena occurring in language use and evaluating the accuracy of those models. By computational model, we mean an abstract description of a process, system or phenomenon that can be implemented as a computer program (i.e., that is computable). Let us explain the meaning of the main terms in the statement of the question and justify their use:

- **computational model of conversational agents:** as we explained above, computational models are abstract descriptions that are computable. In our case, the model will ultimately describe the participants in a conversation (i.e., the conversational agents). This includes rules of expected behaviour for dialogues in the domain, individual goals, conversational obligations, priorities associated with goals and obligations and a dialogue management component.

- **elements:** this refers to the aspects of the model that are either general to most dialogue situations or specific to the domain in which we have focused our research[5]. This means that although we focus part of the research on a specific domain, we expect the conclusions we draw to apply to conversation in general.

- **cooperative and non-cooperative linguistic behaviour in dialogue:** this distinction is central to our research and was hinted at above. Intuitively, it refers to whether participants do or do not behave as is expected for the type of dialogue in which they engage. This will be defined rigorously in Chapter 3 and exhaustively evaluated in Chapter 4.

- **exhibit and cope with:** this means that the agents should be able, not only to produce cooperative and non-cooperative linguistic behaviour, but also to detect it in their interlocutors and eventually reason about it as part of the decision processes that motivate their actions.

- **political interviews:** this is the domain in which we will focus our

[5]For example, the rules of expected behaviour will apply only to dialogues in the domain, while the mechanism by which cooperation is measured and the ability of agents to decide whether they will discharge an obligation or behave following their private goals are properties that would apply to dialogues of any type.

study of non-cooperative conversational behaviour. It is intended to provide a well-defined set of scenarios, scoping the research in a way that is suitable for concrete empirical analysis.

1.4 Why does it Matter?

The motivation for addressing this question is to extend the state-of-the-art of computational dialogue modelling to cases in which the conversation flow is compromised to some extent but without reaching complete breakdown. Shedding light on the nature of linguistic non-cooperation in dialogue promises to yield a better understanding of conversation, and will certainly be of use in the analysis – manual, semi and fully automatic – of natural language interactions and on applications such as human-like virtual personal assistants, tutoring agents, sophisticated dialogue systems, and role-playing virtual humans.

The assumption that dialogue is regarded as an activity that is – or should be – inherently cooperative, has deprived alternative situations of much attention. This is worsened by a lack of clarity in what is considered non-cooperation in dialogue. As discussed in Chapter 2, this has led to studies of conflict and strategic actions in conversation that belong in the realm of non-cooperative behaviour at the level of the social activity, but still assume that participants are cooperative in their linguistic actions.

Looking for an answer to the research question proposed above would shed light on several areas:

- For general knowledge, this research would provide a better understanding of dialogue structure and pragmatics by looking at phenomena that have not been addressed before.

- In the area of dialogue systems, virtual humans and conversational agents, a generalization of our results would allow for the development of systems that deal with non-cooperative conversational behaviour and/or that behave non-cooperatively (according to their goals), resulting in increased flexibility, robustness and closeness to how humans interact using language.

- Direct applications of an accurate model of non-cooperative linguistic behaviour in dialogue to the analysis of natural, everyday conversation.

- For the empirical domain, it would provide a better understanding of the actions of interviewers and politicians during an interview, of the consequences these have on the dynamics of the dialogue and of how all this is perceived by the audience.

1.5 Contributions of the Thesis

This thesis makes the following original contributions to knowledge:

- A definition of cooperative and non-cooperative linguistic behaviour in dialogue, which combines the notions of discourse obligations and dialogue games to specify appropriate behaviour, and allows for the detection of inappropriate actions. The definition is formalised and fully specified for the political interview conversational setting.

- A coding scheme for the manual segmentation, annotation and classification of linguistic behaviour in political interviews. The coding scheme is supported by domain-independent tools, and evaluated for reliability on a corpus of political interviews.

- A domain-independent, automatic method for measuring non-cooperative linguistic behaviour empirically in annotated dialogue. The method is fully implemented, and evaluated for validity on a corpus of political interviews.

- A domain-independent, formal and implementable model of conversational agents that incorporates the concepts and mechanisms above, combining them with the other elements and functions involved in conversation. We show how this addition allows for the generation of a wider range of dialogues, by manipulating parameters that control how agents weigh discourse obligations and private goals when deciding on their contributions.

1.6 Outline of the Thesis

Chapter 2 reviews relevant literature with the aim of connecting our work with previous research. The chapter is structured around the notion of co-operation in dialogue, focusing first on how other authors have addressed the distinction between cooperation at the level of the linguistic action and cooperation at the level of the underlying social activity. Next, the focus is on linguistic cooperation and on how the concepts that allow us to define it have been approached in traditional models of dialogue. This is followed by a discussion on how non-cooperation has been conceived and incorporated in computational theories of dialogue and where this differs from the approach we take in this work. The chapter concludes with a discussion on the nature and analysis of political interviews with the aim of understanding the empirical domain and what should be considered linguistic cooperative behaviour in such a conversational setting.

Chapter 3 presents the conceptual framework for the study of non-cooperation in dialogue used in the rest of the thesis. This includes precise definition of linguistic cooperation and non-cooperation, based on the interconnected notions of global dialogue games and discourse obligations. These concepts are formalised and illustrated in the context of political interviews. The chapter closes with a detailed set of rules that specify expected behaviour in political interviews which is used in the examples of the rest of the thesis and in the empirical evaluation of the approach.

Chapter 4 describes and evaluates a semi-automatic method for measuring linguistic non-cooperation in naturally-occurring political interviews. The method consists of two steps. In the first step, dialogue transcripts are segmented and coded following a proposed scheme for qualitatively classifying the contributions of the speakers. In the second step, the annotated data is automatically analysed with respect to rules that specify expected behaviour in political interviews. The result of this automatic analysis is a dialogue marked up with cooperative and non-cooperative features. These features lead to a score for each speaker that indicates the extent to which the participant behaved according to the expectations associated with their role in the dialogue, which we interpret as the degree of cooperation of the participant with respect to the conversational setting. The chapter includes an extensive evaluation of the reliability of the coding scheme, analysing inter-annotator agreement for segmentation and for both annotation stages on a corpus of political interviews. The validity of the method is assessed by analysing the correlation between the resulting scores and human judgment on the same interview transcripts elicited by means of a survey.

Chapter 5 presents a model of conversational agents that can exhibit and deal with varying degrees of linguistic cooperation. With a focus on dialogue management, the concepts, structures and algorithms introduced and evaluated earlier in the thesis are incorporated in the agents' architecture. The mechanisms by which these elements work together are formalised and illustrated with an example from the political interview conversational setting. It is shown how a simple change of parameters can lead to interactions with very different degrees of linguistic cooperation. The chapter closes with a discussion on the differences and similarities of the modelling approach with related research and with the description of a prototypical system that incorporates some of the elements in the model.

Chapter 6 concludes, with a list of the contributions of the thesis and a discussion of future work.

...we may not be sure whose shoulders we are standing on, but we know whose hands we are holding.

On Whose Shoulders?
Yorick Wilks

Chapter 2

A Review of Related Work

This chapter presents a review of research on dialogue pragmatics and computational dialogue modelling in the light of cooperation. The discussion starts by considering different notions of cooperation from the field of pragmatics. Next, we look at how cooperation and non-cooperation have been addressed in existing approaches to dialogue modelling and analyse in some detail those that are closer to the one put forward in this thesis. Finally, we discuss relevant research in the analysis of political interviews, the domain of our empirical investigations.

2.1 Overview of the Chapter

The notion of cooperation between dialogue participants is recurrent in the analysis and modelling of conversation. This is because conversation requires that participants work together coordinating their actions. However, conversation supports social activities in which participants do not necessarily share the same goals. This often leads to compromises in the extent to which they cooperate and to unexpected phenomena emerging in the dynamics of the interaction. Over the next four sections we look at relevant research in the philosophy of language, pragmatics, computational linguistics, conversation analysis and dialogue systems design. We examine relevant concepts from theories of conversation that lend themselves as starting points to the study of non-cooperative behaviour. We also consider technical aspects of these theories when implemented as part of a conversational agent or of a dialogue system. As we will argue, although many of these approaches deal with non-cooperation at the level of the activity supported by the conversation, they still assume that participants are linguistically cooperative. In contrast, our research is aimed at formalising and studying, systematically, some of the phenomena that emerge when these restrictions are relaxed and participants can be non-cooperative also at the linguistic level.

The empirical domain from where we draw our examples and the data for the corpus study presented in Chapter 4 is that of political interviews. We will briefly present key contributions from sociolinguistics and other disciplines on systematic studies of this dialogue type. In Chapter 3, these insights are developed into a set of rules, capturing behaviour that is conventionally expected from interviewers and interviewees throughout an interview.

Section 2.2 focuses on notions of cooperation in dialogue, reviewing traditional literature and grounding the claim that cooperative and non-cooperative behaviour do indeed happen at different levels of the interaction. Section 2.3 looks at how cooperation is specified and at the role it has played in the analysis and modelling of dialogue. It also discusses the concepts that have been used in the past to model cooperative dialogues and those that we will build on to put together our approach. Section 2.4 focuses on how non-cooperative behaviour has been characterised in the past and on the ways in which these models fail to account for linguistic non-cooperation. Finally, Section 2.5 looks at how political interviews are conventionally characterised, what inappropriate behaviours have been identified on either role and why they provide a rich domain on which to carry out the empirical studies that support this thesis.

2.2 Notions of Cooperation in Dialogue

An assumption that interlocutors fully cooperate is at the core of most of the traditional literature on the pragmatics of dialogue. For instance, Grice's notion of **conversational implicature** provides an explanation for coherence in the following example (Grice, 1975, p. 51):

Example 2.1.

> A is standing by an obviously immobilized car and is approached by B; the following exchange takes place:
>
> (1) A: *I am out of petrol.*
>
> B: *There is a garage round the corner.*

Nothing in B's contribution explicitly indicates that the garage is open at the moment and that it has petrol for sale. However, as A assumes that B is trying to help, he or she can conclude that B thinks it to be the case. This information about the garage is not logically entailed by B's words; it is **implicated**.

For reasoning about such cases, Grice provides a descriptive framework, starting with the **Cooperative Principle** (Grice, 1975, p. 45):

> *Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.*

The principle is then divided into **conversational maxims**: a set of rational principles which Grice grouped following the Kantian categories of **Quantity**, **Quality**, **Relation** and **Manner** (or **Modality**), as shown in Figure 2.1. Grice established a relation between the maxims and the Cooperative Principle (CP): if the maxims are followed, the principle is observed. Conversely, assuming the principle holds, observance or **exploitation** of the maxims, allows the listener to work out what the speaker is really trying to say. This notion is called **conversational implicature**. Despite appearing as imperatives, as has been noted e.g. by Prince (1982), the CP and the maxims are of most use if regarded as presumptions that speakers and listeners can exploit for conveying and inferring meanings that are not logically entailed by their utterances. For example, if we assume the speaker is trying to help, the mention of a garage in Example 2.1 implicates that it is open and selling petrol, as otherwise he or she would be violating the Maxim of Relation. Maxims can also be **exploited**, that is flouted with the

Maxim of Quantity:

1. Make your contribution as informative as is required (for the current purposes of the exchange).
2. Do not make your contribution more informative than is required.

Maxim of Quality: Try to make your contribution one that is true.

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Maxim of Relevance: Be Relevant.
Maxim of Manner: Be perspicuous.

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief.
4. Be orderly.

Figure 2.1: Grice's conversational maxims (Grice, 1975, pp. 45–46)

purpose of inducing an implicature. Consider the following example:

Example 2.2.

(2) A: *How did you find the restaurant last night?*

B: *It was cheap and the toilets were clean.*

B certainly knows, for instance, about the quality of the food. The omission of this information is an apparent violation of the Maxim of Quantity, which implicates that the food in the restaurant was not good.

Furthermore, speakers can also covertly **violate** or overtly **opt out** from the maxims or from the CP. Grice refers to these possibilities, although he does not elaborate on them any further (Grice, 1975, p. 49):

> A participant in a talk exchange may fail to fulfill a maxim in various ways, which include the following:
>
> 1. He may quietly and unostentatiously VIOLATE a maxim; if so, in some cases he will be liable to mislead.
> 2. He may OPT OUT from the operation both of the maxim and of the CP; he may say, indicate or allow it to become plain that he is unwilling to cooperate in the way the maxim requires. He may say, for example, *I cannot say more; my lips are sealed.*

Although extremely influential, Grice's ideas have been often criticised as being too vague, e.g. by Kiefer (1979), Sperber and Wilson (1982) and, more recently, Frederking (1996) and Clark (1996, pp. 141–146); limited by intercultural differences, e.g. by Keenan (1976), later contested by Prince (1982); or applicable only to cases in which there is a strong sense of co-operation between the interlocutors, e.g. by Asher and Lascarides (2008;

in press). Most of these shortcomings are evident if Grice's maxims and the CP are regarded as rules or guidelines that strictly govern the behaviour of participants in a conversation. Grice might have contributed to the misconception, as he states: "For a time, I was attracted by the idea that observance of the CP and the maxims, in a talk exchange, could be thought of as a quasi-contractual matter, with parallels outside the realm of discourse" (Grice, 1975, p. 48).

One consequence of regarding Grice's framework as normative is that very little attention has been paid to cases in which the operation of the CP cannot be assumed[1]. This behaviour, which Grice referred to as **opting out**, does not necessarily lead to a breach of expected or rational behavior. Recall, for instance, the example introduced in the previous chapter of a witness under interrogation in a US trial who refuses to answer a question by appealing to the Fifth Amendment of the Constitution. Prince (1982) presents a convincing analysis of the social role of the maxims in each category, emphasising that they are of most use when taken as presumptions, i.e. as underlying hypotheses that speakers and listeners can exploit for conveying and inferring meanings that are not logically entailed by the utterances. This is also the approach taken by Levinson (1983) when discussing Grice's work.

The cooperative principle refers to the "accepted purpose or direction of the talk exchange". Although vague, this mention of purpose or direction could be taken to mean shared purposes or common goals. In fact, Grice does mention these when he elaborates on the concepts of exchange

[1] An exception is Asher and Lascarides (in press) who propose an extension to previous work (Asher and Lascarides, 2003) that explains the derivation of implicatures even when Gricean cooperation cannot be assumed. The authors use a notion of safety to relate implicatures derived outside the CP with the speaker's public commitments and distinguish between three levels of cooperation: basic, rhetorical and Gricean. We come back to this work and the connections with our approach later.

and engagement as they are used in the statement of the CP. He refers to "cooperative efforts; and each participant recognises in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction" (Grice, 1975, p. 45) and assumes that "each party should, for the time being, identify himself with the transitory conversational interests of the other" (Grice, 1975, p. 48). This contributes to the claim that the notion of cooperation put forward by Grice is strong (Asher and Lascarides, 2003; Asher and Lascarides, in press), that is one in which participants have a common goal and willingly adopt each others intentions. In the rest of this section we see how cooperation has been characterised at levels that are in a sense weaker than proposed by Grice.

Attardo (1997) revisits Gricean pragmatics, identifying two levels of cooperation related, respectively, to linguistic and non-linguistic goals. He claims that Grice's definition of the CP exhibits a "systematic ambiguity" in this regard. Attardo supports his view by showing how some implicatures in the examples given by Grice (1975) are worked out relying on goals that must be shared beyond that particular conversational exchange. In Attardo's view, linguistic cooperation refers to assumptions on the speakers' behaviour in order to encode and decode intended meaning. Non-linguistic cooperation is related to the behaviour of the participants towards realising the goals they intended to achieve by means of the exchange. Attardo goes on by proposing an additional principle of non-linguistic cooperation, which he calls **Perlocutionary Cooperative Principle** (PCP), regarding Grice's CP as relevant only to the linguistic level – an interpretation of Grice's work which is debatable, as it follows from the previous paragraph. Attardo does not claim that the theory based on Grice's CP is wrong, but instead proposes this explicit distinction between linguistic and non-linguistic cooperation as an expansion. The article includes an analysis of the relation between the PCP and other "principles" such as politeness or self-interest, which in Attardo's view override the PCP. Unfortunately, his discussion on the violation of the PCP is inconclusive and limited to observing the interdependence of the CP and PCP for working out implicatures, which follows directly from the fact that Grice intended his CP to account for both linguistic and (at least at some level of the interaction) non-linguistic cooperation. Attardo is sceptic about the possibility of meaningfully violating or flouting the PCP, somewhat close to Grice's initial temptation to regard the CP as a quasi-contractual matter. This stance seems to suggest that violating the PCP would be irrational and therefore unworthy of attention. From such a normative perspective, the PCP would suffer from the same criticisms and shortcomings that were identified in relation to the normative interpretation of Grice's CP.

Asher and Lascarides (2008) also identify two levels of cooperation in Grice's approach: a first level related to how meaning is encoded by the participants (Clark, 1996) and a second level in which the CP and the conversational maxims are followed. In recent work (Asher and Lascarides, in press), they explicitly address the problem of coherence in strategic conversation – e.g. courtroom cross-examinations – in the absence of full cooperation between the participants. In this case, they distinguish between three levels of cooperation: **basic** and **rhetorical cooperation** governing respectively linguistic meaning and speech act coherence, and **full** or **Gricean cooperation** relating to the alignment of conversational goals. Assuming basic cooperation, the authors focus on rhetorical cooperation and provide a formal mechanism for drawing implicatures even when the CP does not hold. They propose a property of **safety** to determine when implicatures conveyed in settings in which Gricean cooperation cannot be assumed can be taken as public commitments. Asher and Lascarides refer to plans for empirical work to extend their approach beyond the naturally-occurring, yet isolated examples they use in the article (Asher and Lascarides, in press, p. 49). The first steps in these directions are reported in the context of the STAC Project (Asher et al., 2012; Afantenos et al., 2012).

A different approach to distinguishing between goals in conversation is proposed by Walton and Krabbe (1995). In the context of natural argumentation (or informal logic), they develop a formal, normative framework aimed at identifying situations that lead to **fallacy**[2]. In doing so, they define a typology of dialogue given by the initial situation (precondition) that triggered the exchange, the joint aims assumed to be shared by the participants and their individual goals (which can be at odds). Walton and Krabbe identify six main types of dialogue in their framework: persuasion, negotiation, inquiry, deliberation, information-seeking and eristic[3] – Figure 2.2 shows their characterization of information-seeking dialogues. In the course of a conversation, participants usually reach situations in which the preconditions for a dialogue of a different type are met. The change from one dialogue type into another is called **dialectical shift** and must be acknowledged by both parties. When the second dialogue appears as a sub-dialogue of the first (i.e. when the second dialogue ends the first one resumes from the point where it was left) the shift is called a **functional embedding**.

Walton and Krabbe addressed the problem of formally modelling persuasion dialogue from a game-theoretic perspective. They limited their study to persuasion dialogue as it was directly relevant to natural argumentation.

[2]In informal logic, a fallacy is an argument that appears as valid but has flaws in the reasoning from the premises to the conclusion.

[3]Eristic dialogues are essentially verbal fights; arguing for the sake of conflict.

Information-Seeking Dialogue

Initial Situation	Personal Ignorance
Main Goal	Spreading Knowledge & Revealing Positions
Participants' Aims	Gain, Pass on, Show, or Hide Personal Knowledge
Side Benefits	Agreement, Develop Reveal Positions, Influence Onlookers, Add to Prestige, Vent Emotions
Subtypes	Expert Consultation, Didactic Dialogue, Interview, Interrogation

Figure 2.2: Characterisation of information-seeking dialogues (Walton and Krabbe, 1995, p. 66)

Building on an example (presumably artificial, given the absence of sources), they identified a set of strict rules for two subtypes of persuasion dialogue (permissive and rigorous) and defined a third type as the functional embedding of a rigorous persuasion dialogue within a permissive persuasion dialogue. The dialectic shifts at the transitions between both types were also modelled by a strict set of rules, defining what was a **licit shift** and what was not. The authors claim that their model is general enough to account for dialogues like the example and their thesis is that if both participants follow the rules and all shifts are licit, the argument is valid. On the other hand, failure to follow the rules or to perform licit shifts results in fallacy.

The limitations of this approach are mainly related to the lack of empirical support. Walton and Krabbe accept that a formal account can only deal with an idealization of the type of dialogue it addresses. Nevertheless, they consider the attempt to provide a mathematically precise description of (a certain class of) dialogue to be worthwhile, even if it does not apply to the vast complexity of naturally occurring conversation. We agree with this view, which is especially appealing if we attempt to approach dialogue phenomena from a computational perspective, but reckon that resulting models would be greatly improved if they were supported (and inspired) by stronger empirical evidence.

An analysis of Walton and Krabbe's framework from the perspective of cooperation appears in an article by Reed and Long (1997). The authors propose a definition of cooperation in dialogue that, they claim, acts at the level of the discourse, as opposed to the utterance-by-utterance notion of cooperation in Grice's CP. In their view, participants are cooperative if they follow a common set of dialogue rules, and stay within a mutually acknowledged framework. This means that interlocutors agree on the type of dialogue (in Walton and Krabbe's sense) in which they engage, adhere to the set of rules and main goal of that type and respect the rules for licit functional embedding when shifting from one type of dialogue to another. The authors use the term **collaboration**[4] to refer to a notion stronger than cooperation in which the individual goals of both participants point in the same direction. They continue with an analysis of Walton and Krabbe's typology in terms of cooperation and collaboration and conclude that all dialogue falling in the classification is inherently cooperative. This is consistent with Walton and Krabbe's conclusion that, for natural argumentation, failure to follow the rules they identified for persuasion dialogue results in fallacy. However, it is not clear what are the consequences of breaking the rules for other types of dialogue and Reed and Long do not address this issue in their article.

[4]Although we adopt the distinction between what Read and Long call cooperation and collaboration, in the sequel we will abandon these terms in favour of the more explicit (and less arbitrary) **linguistic cooperation** and **task-level cooperation**. Coincidentally, the term collaboration is used by Allwood et al. (2000) to denote a notion weaker than that of **ideal cooperation** (Allwood, 1976) which involves the participants cognitive and ethical mutual consideration, joint purposes and trust. This is a use of the words cooperation and collaboration with a meaning somewhat opposite to that intended by Reed and Long.

We agree with the authors in that their definition of cooperation operates at a more global level in the discourse than that of Grice. Furthermore, it allows for an analysis of a participant's behaviour from a perspective that includes the obligations or presumptions determined by the social setting, the individual goals of each party and the benefits they expect from the exchange – that is, the same aspects Walton and Krabbe considered for establishing their typology of dialogues[5].

From the discussion so far, it is clear that cooperation in dialogue can be addressed from several perspectives. The notion of expected behaviour is recurrent in the literature. Determined by low-level conversational mechanisms, social convention, institutional settings or participant roles, underlying assumptions on speakers behaviour are relevant to the resulting structure of the conversation. This is also evident in the short examples of non-cooperative behaviour we presented in the previous chapter. Consider for instance the following turns from the interview in Example 1.1 in Chapter 1:

(27)	PAXMAN:	I note that you're not answering the question whether you threatened to overrule him.
(...)		
(32)	HOWARD:	You can put the question and I will give you an answer.

and this fragment from the interview in Example 1.3:

[5] As discussed in the next section, Bunt (1994) also looks at these aspects when defining how dialogue acts operate on context. He uses the term **social context** to refer to the institutional setting of the dialogue, the roles of the participants and their communicative rights and obligations at any point in the dialogue. Individual goals are also part of the account and fall under what Bunt calls **cognitive context**.

(30) GALLOWAY: You are actually conducting one of the most - even by your standards - one of the most absurd interviews I have ever participated in.

In both cases, participants appeal to underlying assumptions on the expected behaviour associated with their respective roles in a political interview. When Galloway classifies the current interview as absurd, he does so by contrasting it with other instances, certainly closer to the convention. Likewise, when Paxman notes that Howard is not answering the question, he does so on the basis that interviewees are expected to give an answer, as in fact Howard confirms shortly after. These aspects of political interviews are discussed in detail in Section 2.5.

We close this section on two notes. First, Gricean maxims are of better use when regarded as underlying assumptions that participants can utilise to increase the amount of information conveyed by their utterances. Analogously, awareness of expected behaviour determined by the type of conversation – and awareness that those expectations can be disregarded – offer an interesting perspective from which dialogue can be analysed and modelled. Second, cooperation can be observed at several levels. At each of these levels, a set of (usually tacit) rules, norms or conventions is followed by cooperative participants when producing and interpreting their contributions. Two very distinctive levels can be identified in the research discussed above: **linguistic cooperation**[6], in which participants adhere to a set of linguistic conventions and mutual expectations specific to the conversational setting

[6] Asher and Lascarides (in press) further divide this level into **basic cooperation** (Clark, 1996) which relates to how meaning is encoded (e.g. whether the participants speak the same language, or use specific jargon with the same meaning) and **rhetorical cooperation** which refers to what dialogue acts are expected in response to the acts of the others party. In these terms, we will assume that basic cooperation is the case and focus on a notion of cooperation closer to Asher and Lascarides's rhetorical cooperation.

that preserve the dynamics of the conversation; and **task-level cooperation**, in which the individual goals of the participants with respect to the underlying task are aligned.

2.3 Cooperation in Dialogue Modelling

As we mentioned in the previous chapter, most computational models of dialogue are constructed on the assumption that participants are cooperative. These assumptions are reasonable, as such models are developed in the context of **dialogue systems**: software systems that have a natural language dialogue component as part of their interface.

Dialogue systems are built for the purpose of providing a service to their users. In this scenario, failure to cooperate, either on the side of the system or of the user, is against the premises on which the system is conceived and used. Examples of early academic dialogue systems include ARTEMIS (Sadek et al., 1997), VERBMOBIL (Wahlster, 1993) and TRAINS (Allen and Schubert, 1991). ARTEMIS is a spoken information-seeking system based on principles of rationality (Cohen and Levesque, 1990) and intentionality (Sadek, 1992; Sadek et al., 1996). VERBMOBIL is a speech-to-speech translation system that works as a mediator between two users speaking different languages, relying on dialogue acts (Bunt, 1994) and plan recognition (Jekat et al., 1995; Alexandersson et al., 1997).

TRAINS is a planning system for the transportation domain, which was influential in developing and trying new ideas for dialogue modelling. We describe the system in some detail below, as the theoretical approach to dialogue management is closely related to ours. Before, we discuss two concepts that are relevant to cooperation and dialogue modelling: **dialogue**

acts (Bunt, 1994) and **discourse obligations** (Traum and Allen, 1994)[7].

In the speech act tradition (Austin, 1962; Searle, 1969; Searle, 1979), dialogue is structured as a sequence of actions performed by the speaker with associated **force** and **propositional content**. **Dialogue acts**[8] (Bunt, 1994) are an extension of speech acts to include features from turn-taking, adjacency pairs and grounding. Originating from the field of **conversation analysis** (Schegloff, 1968), it was first proposed by Sacks et al. (1974) that dialogue participants take turns for making their contribution at adequate places and in particular ways. These shifts are governed by a set of **turn-taking** rules. **Adjacency pairs**, also originate from conversation analysis and refer to conversational structures composed of two parts produced by different speakers (e.g. question-answer, greeting-greeting, offer-acceptance/rejection). **Grounding**, or establishing a **common ground**, refers to the process by which participants agree on the set of things they mutually believe, know or assume (Clark and Schaefer, 1989). Grounding is divided in two phases (Clark and Brennan, 1991, p. 130):

> **Presentation phase:** A presents utterance u for B to consider. He does so on the assumption that, if B gives evidence e or stronger, he can believe that she understands what he means by u.
>
> **Acceptance phase:** B accepts utterance u by giving evidence e that she believes she understands what A means by u. She does so on the assumption that, once A registers that evidence, he will also believe that she understands.

[7] Although they are not part of this review, other influential notions like cognitive states (Allen and Perrault, 1980; Allen, 1995) and plans (Grosz and Sidner, 1990; Chu-Carroll and Carberry, 1998) are also present in the TRAINS system.

[8] Dialogue acts are closely related to **conversational moves** (Power, 1979; Carletta et al., 1997) and **conversation acts** (Traum and Hinkelman, 1992).

Traum (1994) revisited the notion of grounding and proposed a computational theory of the mechanisms involved in achieving a common ground in natural dialogue.

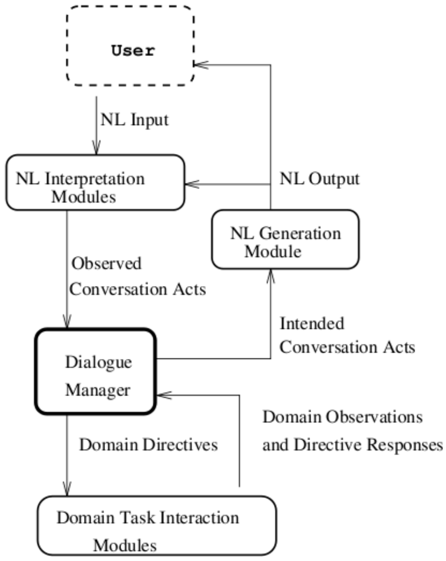
Bunt proposes dialogue acts as "functional units used by the speaker to change the context" (Bunt, 1994, p. 3) and identifies in them three properties: **utterance form**, **communicative function** and **semantic content**. Changes to the linguistic context are determined by the utterance form, while communicative function and semantic content relate to the force and propositional content of a speech act, respectively. From a context-changing perspective, the communicative function determines the significance of the semantic content in the new context. Bunt gives the following example (Bunt, 1994, p. 4):

> For instance, a dialogue act with the utterance form "*Does it rain?*", the communicative function YES/NO QUESTION and the proposition `it is raining` as semantic content, has the effect of adding the utterance *Does it rain?* to the linguistic context, and creating in the addressee (among other things) the belief that the speaker wants to know whether the proposition `it is raining` is true.

The notion of context proposed by Bunt considers five categories: **linguistic**, **semantic**, **physical**, **social** and **cognitive**. Each category is further divided into **global** and **local**. Global aspects for each category remain constant during the conversation but local context changes as the dialogue progresses. The framework is completed by distinguishing between **dialogue control** and **task-oriented** dialogue acts, depending on whether their communicative function is intended to control the interaction or concerned with the underlying task.

Discourse obligations (Traum and Allen, 1994) were introduced as an alternative to joint intentions (Cohen and Levesque, 1991) and shared plans (Grosz and Sidner, 1990) to allow for models of dialogue in which participants do not have the same high-level (i.e. non-linguistic, task-level) goals but still engage in conversation. In this view, conversational behaviour is determined, not only by participants' goals, but also by "a sense of obligation to behave within limits set by the society that the agent is part of." (Traum and Allen, 1994, p. 2). Obligations are obtained from rules that encode discourse conventions and updated dynamically along the course of the conversation. In the case of conflict between goals and obligations, the latter are favoured. The authors consider the possibility of an agent pursuing its goals at the expense of violating obligations, but they do not analyse the consequences of this any further. In fact, as we discuss below, the implemented system on which they tried these ideas always discharges obligations before considering any private goals.

We return now to the TRAINS dialogue system (Allen and Schubert, 1991; Traum and Allen, 1994), and more specifically to the dialogue manager component, from the perspective of cooperation. As usual in most dialogue systems, the **dialogue manager** (see Figure 2.3) controls the structure of the conversation towards a high-level, task-related goal. In the case of TRAINS this amounts to obtaining a transportation plan according to the requirements of the user. The dialogue manager connects the dialogue acts – or conversation acts, as they are called in TRAINS – that result from interpreting the user's utterances with domain-specific tasks modules. It then decides what dialogue acts to perform based on the outcome of these modules, and sends them to the natural language generation module that produces the utterances as the system's next move. This is specified by the algorithm in



(Traum and Allen, 1994, p. 4)

Figure 2.3: Architecture of the TRAINS Dialogue System

Figure 2.4. The priorities for goals, intended contributions and obligations are in the following (fixed) order:

1. Discourse obligations associated with adjacency pairs
2. Weak obligation: don't interrupt the user
3. Intended contribution
4. Weak obligation: grounding
5. Discourse goals: plan negotiation
6. High-level discourse goals: form a shared plan

From these priorities and the algorithm it is clear that the conversational behavior of the system is cooperative in the sense that obligations will be discharged before any private goals are considered. This is consistent with the purpose of the system, designed to assume the user's private goals as

```
(1)  while conversation is not finished
(2)      if system has obligations
(3)      then address obligations
(4)      else if system has turn
(5)      then if system has intended conversation acts
(6)           then call generator to produce NL utterances
(7)           else if some material is ungrounded
(8)           then address grounding situation
(9)           else if some proposal is not accepted
(10)          then consider proposals
(11)          else if high-level goals are unsatisfied
(12)          then address goals
(13)          else release turn
                  or attempt to end conversation
(14)     else if no one has turn
(15)     then take turn
(16)     else if long pause
(17)     then take turn
```

(Traum and Allen, 1994, p. 5)

Figure 2.4: Discourse Actor Algorithm of the TRAINS Dialogue System

its own[9]. Obligations are addressed much earlier than private goals are considered. However, the emphasis on obligations turns the approach into a powerful mechanism for addressing issues unaccounted for by other models. In the rest of the section we look at a line of research that emerged from the introduction by Traum and Allen (1994) of discourse obligations, as it provides a suitable context in which to frame our contribution.

2.3.1 Discourse Obligations

By proposing to focus on discourse obligations for explaining the actions of speakers in conversation, Traum and Allen (1994) initiated a line of research developed further by Poesio and Traum (1997; 1998), Matheson et

[9]This is similar to intention-based approaches to cooperation in which systems assume the user's intentions as their own and then reason to act following those intentions. Galliers (Galliers, 1988) calls this kind of unconditional cooperation **benevolence**.

al. (2000) and Kreutel and Matheson (1999; 2000; 2001; 2003b). The model of conversational agents presented in Chapter 5 draws elements from these works, which we summarize below.

Poesio and Traum Theory (PTT)

PTT was first proposed by Poesio and Traum (1997; 1998) on the basis that participants' actions in a conversation be part of the common ground, with the aim of providing a unified treatment for discourse context in reference resolution, intention recognition and dialogue management. The authors claim that including explicitly the occurrence of conversation acts, as opposed to just their (domain-dependent) propositional content, as part of the dialogue situation that is agreed upon between participants, represents a shift from modelling the meaning of contributions to modelling their use. This, in turn, allows considering pragmatic information in the deliberation process for managing the dialogue.

The dialogue acts in PTT are taken from Conversation Act (CA) theory (Traum and Hinkelman, 1992): an extension of the theory of speech acts, to consider aspects from turn-taking, adjacency pairs and grounding. In speech act theory, a locutionary act (i.e. an utterance) usually generates several illocutionary acts (e.g. inform, request, accept). These are called **core speech acts** in CA and are required to be grounded before taking full effect. Grounding is achieved by means of **grounding acts** (e.g. acknowledge, request repair). CA also considers **turn-taking** (e.g. take turn, keep-turn, release-turn) and **argumentation acts** (e.g. elaborate, clarify). The latter are complex, domain-dependent acts, that can take whole conversations to complete and could be regarded as the dialogue games as we use them in our thesis (see section on global dialogue games below). The following table lists the four types in increasing order of complexity[10](Poesio and Traum, 1997):

Discourse Level	Act Type	Example
Sub Utterance Unit	Turn-taking	take-turn, keep turn, release-turn
Utterance Unit	Grounding	initiate, continue, acknowledge, repair, request-repair, request-acknowledge
Discourse Unit	Core Speech Act	inform, yes-no-question, evaluate, suggest, request, accept, reject
Multiple DUs	Argumentation Acts	elaborate, summarize, clarity, question-and-answer, convince

To represent the occurrence of conversation acts in the common ground, Poesio and Traum use a reinterpretation of Muskens' compositional DRT (Muskens, 1994), where DRSs representing the propositional content of conversation acts are seen as transitions between dialogue situations.

Poesio and Traum (1998) continue their proposal on the treatment of conversation acts in the representation and dynamics of the dialogue situation (now called conversational score), with emphasis on aspects not studied in previous works on speech acts: grounding and participants' obligations. The paper focuses on core speech acts and grounding acts, using a task-independent dialogue act taxonomy based on the Discourse Resource Initiative (1997) (DRI):

- Locutionary acts, representing the utterances of speakers.
- Core speech acts, further classified in terms of their effects on social attitudes (e.g. obligations) to be addressed in the conversation:
  - FORWARD-LOOKING FUNCTION: introduce new social commitments (obligations). Obligations can be imposed on the speaker (e.g. commit) or on the hearer (e.g. info-request) and can be conditional (e.g. offer-accept).

[10]Utterance Units are intonation phrases, i.e. the spoken analogue to written sentences.

  - BACKWARD-LOOKING FUNCTION: a response to previous acts (e.g. accept, answer). Usually discharge obligations, but can also introduce new ones (e.g. answer introduces an obligation for the hearer to acknowledge the answer).

- Grounding acts: have the effect of moving information from ungrounded to grounded sections of the conversational score.

Each dialogue participant keeps a **conversational score**: a record of the dialogue acts, as well as public beliefs, intentions and social commitments of each participant, with grounded and ungrounded information. Each act leads to an update of the conversational score with one or more discourse units. Discourse units are are first added to the score as ungrounded units. Grounding acts cause a transfer of these units to the grounded section of the score. Forward-looking acts typically introduce discourse obligations once they are grounded. In contrast, appropriate backward-looking acts discharge obligations. These effects are specified by means of **update rules** with preconditions on the current state of the conversational score and effects representing the changes.

Matheson et al. (2000) implemented part of PTT in the context of the TRINDI project using the TrindiKit dialogue engine (Larsson et al., 2000; Larsson and Traum, 2000) and information states update rules (Traum and Larsson, 2003). The theory was implemented as the EDIS dialogue system with a focus on dialogue management, more specifically, on aspects of grounding and on the management of obligations. Discourse obligations are tied to dialogue acts by means of update rules. Obligations are introduced by applying these rules when the specified dialogue acts are grounded in information states that satisfy their preconditions. An obligation is discharged when the dialogue act it refers to appears, grounded, on the dialogue history. The system uses intentions to represent speaker goals, but maximum cooperation is assumed: the system always chooses to meet its obligations before paying attention to its goals. This accounts for behaviours like those that initially motivated Traum and Allen's (1994) introduction of discourse obligations: e.g. participants responding at all to questions they do not wish to answer. This policy also means that non-cooperative behaviour is precluded, as participants must act against their wishes whenever their obligations are in conflict with their goals.

Obligation-Driven Dialogue Modelling

The modelling approach proposed in this thesis is closest to that taken by Kreutel and Matheson (1999; 2000; 2003a; 2003b). In a series of papers and building on Poesio and Traum's theory, the authors focus their attention on discourse obligations and show how several of the structures used as primitives in other approaches, such as Grosz and Sidner's (1986) intentions, Carletta et al.'s (1997) conversational games and Ginzburg's (1996; 1997) question under discussion, can be derived from a model based on obligations. In addition, the approach is able to account for phenomena left unexplained by these approaches or that require controversial assumptions, such as Boella et al.'s (1999) proposition that cooperation results form the assumption that participants aim at all times to avoid offending each other so that the dialogue progresses smoothly.

Kreutel and Matheson (2001) explore the suitability of obligation-driven modelling for explaining strategic behaviour in discussion scenarios, as opposed to cooperative behaviour in which the participants assume each other's intentions to motivate their own actions. The authors illustrate their model with a few handcrafted examples and do not attempt an empirical study as the one we present in Chapter 4.

Kreutel and Matheson regard conversational cooperation essentially in the same way we do in this thesis: participants are conversationally cooperative if they act according to their obligations. They refer to cooperation at the level of intentions as traditional cooperation and this is where their model accounts for strategic acting. In other words, participants in their model are conversationally cooperative, contrary to where the focus of this thesis rests. We will come back to this in the Section 2.4.

2.3.2 Global Dialogue Games

A notion central to our conception of cooperation is that of a **dialogue game**. The term has been used extensively in the literature to refer to different concepts. In this thesis, a dialogue game will denote a set of rules that apply to entire conversations. When needed, we will use the phrase **global dialogue game** to emphasise the difference between our notion and others with similar names in the literature.

Global dialogue games share some characteristics with Levinson's (1979, p. 368) notion of **activity type**:

> "it refers to any culturally recognized activity, whether or not that activity is co-extensive with a period of speech or indeed whether any talk takes place in it at all. In particular I take the notion of an activity type to refer to a fuzzy category whose focal members are goal-defined, socially constituted, bounded, events with *constraints* on participants, setting, and so on, but above all on the kinds of allowable contributions. Paradigm examples would be teaching, a job interview, a jural interrogation, a football game, a task in a workshop, a dinner party and so on."

The main difference would be that, while Levinson considers an activity type as a "fuzzy category", our dialogue games are precisely defined by means of formal rules, as exemplified towards the end of Chapter 3. These rules are then used as the basis for the definition of cooperative conversational behaviour and also specify the effects of the participants' contributions. This is consistent with Levinson's conclusions that activity types play a central role in language usage in two special ways: constraining what actions are considered allowable contributions to each activity, and helping determine how contributions are to be interpreted (Levinson, 1979, p. 393).

Global dialogue games are closer to the notion Walton and Krabbe (1995) utilise in their study of informal logic and argumentation dialogues, building on the **dialectical systems** Hamblin (1970) used for studying fallacies. As mentioned earlier in the chapter, Walton and Krabbe propose a typology of dialogue types by specifying elements such as the participants, the initial and final dialogue situations, the actions available to the participants and, among these, those that are suitable for the role of each participant at specific points in the dialogue, etc. Several authors have built on their approach to model other types of conversation, including human-computer debate (Maudet and Moore, 2001), computer-mediated crosslingual communication (Piwek et al., 2007) and purchase negotiation, argumentation and conflict resolution in agent communication for multi-agent systems (McBurney and Parsons, 2009; Karunatillake et al., 2009).

Piwek (1998, Chapter 6) also builds on Hamblin's dialectical systems and draws on insights from discourse and conversation analysis to model a number of naturally occurring conversational structures in terms of a conversational game. Piwek's games are divided into **conversational stores** that contain dynamic informational elements of the conversation, and **con-**

versational rules that constrain the timing and content of the participants' contributions and how these affect the conversational store. Analogous to the use of discourse obligations proposed by Traum and Allen (1994), Piwek shows how commitments in the conversation store can be used to account for the pressures on the dialogue participants to react in specific ways to assertions, imperatives and questions, as an alternative to less flexible, stack-based models. A similar approach is reported by Beun (2001) for enabling the generation of coherent elementary dialogues.

Ginzburg (2012, Chapter 4) describes a notion similar to our dialogue games to characterise coherent dialogues and states that "the dialogue analyst describes conventionally acceptable moves and the effects they give rise to among conversation participants in terms of information states." (Ginzburg, 2012, p. 61). A dialogue is coherent in this context if there is a sequence of information states that results from applying these rules to the participants' contributions. As our focus is precisely on dialogues in which rules are not followed, it is not clear how Ginzburg's approach would deal with such cases which, by his definition, would be regarded as incoherent.

As said earlier, global conversational games are not to be confused with what in the literature has often been referred to as dialogue games or similar terms (Levin and Moore, 1977; Carlson, 1983; Mann, 1988; Carletta et al., 1997; Pulman, 2002). These **local dialogue games** are extensions of speech acts to consider goals and span short exchanges, also called **conversational procedures** (Power, 1979), **conversational games** (Kowtko et al., 1992; Pulman, 1997; Lewin, 2000) and **dialogue macrogames** (Mann, 2002).

2.4 Previous Approaches to Non-Cooperation

There have been previous approaches to modelling dialogue on the basis that participants are not always fully cooperative. The first approach we will consider belongs to an area of research in which the construction of dialogue models relies strongly on a detailed description of the user, i.e. a **user model** (Wahlster and Kobsa, 1986). With the aim of analysing a user's perception of the user model in a dialogue system and the speaker's strategies for maintaining a certain image in the listener, Jameson (1989) presents an extensive study for modelling bias, individual goals, projected image and belief ascription in conversation.

The study is realised by simulating a series of increasingly less cooperative situations in the domain of job interviews, where participants are expected to project a certain image, hide biased opinions, etc. For each utterance, the speaker selects a certain **comment**, by computing the expected **impression** it will make in the hearer, with respect to the image the speaker is trying to project. For anticipating the effect of a comment in the hearer (i.e. the impression), Jameson departs from Gricean pragmatics and proposes **pragmatic interpretations** as a basis for reasoning. He claims this mechanism to be more general than implicatures, as it does not rely on a notion of cooperation. Pragmatic interpretations operate as follows:

- Possibility p is not ruled out by comment c (possibly silence);
- but then, if p were true, comment c' would have been made instead of c, since it would have had a more desirable impact on the listener's impression;
- therefore p is apparently not realized.

Jameson implemented some of these ideas, in the context of used cars sales, by means of a dialogue system that can assume different roles (Jameson et al., 1994).

These contributions show that user-model approaches to dialogue modelling are flexible enough to account for situations of an arbitrary degree of intricacy. However, as noted, e.g. by Taylor et al. (1996) the level of detail required in the characterisation of the user and the complexity of mechanism for reasoning about user models can lead to problems like infinite regress in nested beliefs (speaker's beliefs about the hearer's beliefs about the speaker's beliefs...). In the same article, Taylor et al. show that nested beliefs are not necessary when participants are assumed to cooperate in the conversation, if cooperation is restricted to the absence of deception. Taylor (1994) addressed non-cooperative dialogue behaviour by implementing CYNIC, a dialogue system able to generate and recognise deception using a reasoning mechanism equivalent to a theorem prover. A notion of non-cooperation limited to deception is weaker than the one we address in this research.

More recently, Traum (2008) brought attention to the need for computational accounts of dialogue situations in which a broader notion of cooperation is not assumed. As possible applications of such models, he lists intelligent tutoring systems, bargaining agents, personal assistants acting on behalf of their owners and role-playing training agents. These applications have in common that systems cannot afford to adopt the goals of their users as it happens, for instance, in the dialogue systems mentioned above. Traum also provides a list of "behaviours of interest" (along the lines of the non-cooperative features we identified in Chapter 1): unilateral topic shifts or topic maintenance, avoidance, competition, unhelpful criticism, withholding of information, lying and deception, antagonism, etc.

Traum's work on non-cooperative dialogue is mainly aimed at creating virtual humans – or embodied conversational agents (Cassell, 2001) – with abilities to engage in adversarial dialogue. Traum et al. (2005; 2008a) present a model of conversation strategies for negotiation, implemented as a virtual human that can be used for teaching negotiation skills. A recent version of the system (Plüss et al., 2011; Traum, 2012) supports cooperative, neutral and deceptive behaviour, and also is able to reason in terms of secrecy in order to avoid volunteering certain pieces of information. Yet another model, with applications in the domain of training in tactical questioning, is presented by Traum et al. (2007) and Roque and Traum (2007). It can engage in dialogues like the following:

Trainee	Hello Hassan
Hassan	Hello
Trainee	How are you doing?
Hassan	Well, under the circumstances we are fine
Trainee	I'd like to talk about the marketplace
Hassan	I hope you do not expect me to tell you anything
Trainee	I just want to know why people aren't using the marketplace
Hassan	I don't feel like answering that question
Trainee	I think you know something about a tax
Hassan	I am simply doing business. It is rude of you to imply otherwise

(Traum et al., 2007, p. 72)

Both models include variables representing trust, politeness and emotions, and a set of strategies which are selected depending on the values of those variables. These components were developed based on studies of the respective domains and are therefore restricted to them (Roque and Traum, 2007, p. 38).

It must be noted that, despite being adversarial in nature, the kind of conversational scenarios studied by Traum et al. are modelled by means of rules, i.e. the strategies for negotiation and tactical questioning identified by the authors. These rules are followed by the interlocutors, in accordance with the values of certain variables. This means that the dialogues accounted for by these models are adversarial but cooperative under our characterisation of non-cooperative dialogue. From this perspective, it is not clear how effective these models are to account for cases in which participants fail to follow their strategies; an issue at the core of the investigation we present here.

Along similar lines, the work of Kreutel and Matheson (2001; 2003b) described in the previous section accounts for non-cooperative behaviour at the level of the task, what the authors call **strategic acting**. At the conversational level, however, their models – as well as those of Traum and Allen (1994) and Matheson et al. (2000) – always discharge a speaker's obligations before considering their private goals. In our research, we remove this limitation and look at ways in which the resulting phenomena can be characterised, analysed, measured and incorporated into models of conversational agents.

2.5 Analysis of Political Interviews

Before moving on to introducing the conceptual framework we will use in the rest of the thesis, let us discuss key research in the analysis of political interview dialogue: the domain of our empirical studies. Below, we review relevant literature on the analysis of this type of dialogue, discussing early insights from sociology, conversation analysis and more recent ones from social psychology and political science.

As Heritage (1985), Greatbatch (1986; 1988) and Clayman (1988) independently point out, the use of interviews as a means for producing news content has changed notoriously since the outset of broadcasting. In the 1970s interviewers went from a mostly deferential style in which interviews functioned as venues for interviewees (e.g. politicians) to present information (e.g. party policy) unchallenged, to more adversarial exchanges in which journalists would set specific lines of questioning and press for meaningful answers.

This change brought notoriety to the news interviews as a distinct conversational genre and attracted the attention of researchers from the social sciences. Conversation analysts have classified news interviews extensively as an instance of institutional talk (Heritage and Greatbatch, 1991; Clayman and Heritage, 2002). This is a larger class of conversations in which the actions of the participants are distinctively restricted by their institutional roles. As such, participants in news interviews restrict their behaviour according to a pre-allocated subset of possible actions: interviewers ask questions and interviewees respond to those questions. As Greatbatch (1988, p. 404) notes:

> These constraints on the production of types of turns operate with respect to the institutional identities interviewer/interviewee and specify that the incumbents of these roles should confine themselves to asking questions and providing answers, respectively.

Following this, the author identifies some ramifications of these constraints (Greatbatch, 1988, p. 404):

1. IRs and IEs systematically confine themselves to producing turns that are at least minimally recognisable as questions and answers, respectively.

2. IRs systematically withhold a range of responses that are routinely produced by questioners in mundane conversation.
3. a. Although IRs regularly produce statement turn components, these are normally issued prior to the production of questioning turn components.
   b. IEs routinely treat IRs' statement turn components as preliminaries to questioning turn components.
4. The allocation of turns in multiparty interviews is ordinarily managed by IRs.
5. Interviews are overwhelmingly opened by IRs.
6. Interviews are customarily closed by IRs.
7. Departures from the standard question-answer format are frequently attended to as accountable and are characteristically repaired.

Heritage and Greatbatch (1991) and Clayman and Heritage (2002) report on analogous sets of conventions that emerge from performing conversation analysis of news interviews. They ascribe a normative character to these conventions, supported by the institutional setting and the representational roles of the participants: interviewers represent the public and/or the broadcasting company, while politicians represent their parties and/or governmental offices. These representative roles in turn translate, respectively, into requirements for neutrality and into accountability (Heritage and Greatbatch, 1991, p. 96):

> These institutionalized reductions and specialization of the available set of conversational options are, it should be stressed, con-

> ventional in character. They are culturally variable; they are sometimes subject to legal constraints; they are always vulnerable to processes of social change; they are discursively justifiable and are often justified by reference to considerations of task, efficiency, fairness, and so on in ways that the practices making up the conversational "bedrock" manifestly are not. Associated with these various institutional conventions are different participation frameworks (Goffman, 1979), with their associated rights and obligations, different footings and different patternings of opportunities to initiate and sanction interactional activities.

As Schegloff (1988) notes while analysing an infamous exchange between CBS Evening News anchor Dan Rather and then American vice-president George H. Bush during the coverage of the 1988 presidential campaign, participants do not necessarily act always within the pre-established normative institutional framework. A specific type of conversation is achieved through the interaction rather than predetermined by the context. In extreme cases, an exchange that started as an interview can become a conversation of a different type, such as a confrontation in the case of the Rather-Bush encounter[11].

In the following quote – already presented on page 5 of the Introduction –, Heritage refers to some of the actions by which interviewers and interviewees can depart from the convention[12] (Heritage, 1998, p. 8):

[11] A transcript of the Bush-Rather exchange is available at `http://www.ratherbiased.com/transcript.htm` and a video at `http://www.youtube.com/watch?v=FqwQw3THRvU` (last accessed: September 2013).

[12] Heritage's analysis uses a broad notion of turn-taking. Instead, we will limit the use of the term to refer to the system of rules which governs how interlocutors take the floor in a conversation (Sacks et al., 1974). The *actions* mentioned by Heritage above will be regarded as speech acts.

> "the participants -IRs [=interviewers] and IEs [=interviewees]- exclude themselves from a wide variety of actions that they are normally free to do in the give and take of ordinary conversation. If IRs restrict themselves to asking questions, then they cannot – at least overtly – express opinions, or argue with, debate or criticize the interviewees' positions nor, conversely, agree with, support or defend them. Correspondingly, if IEs restrict themselves to answers (or responses) to questions, then they cannot ask questions (of IRs or other IEs), nor make unsolicited comments on previous remarks, initiate changes of topic, or divert the discussion into criticisms of the IR or the broadcasting organization."

After this passage, the author holds that in practice these restrictions are occasionally not observed, but only as departures from the expected behaviour and often resulting in problematic and even sanctionable courses of action. Deviations are thus **sanctionable** and participants are aware of this. Conversational analysts use examples of deviations and their consequences to support the claim that the news interview dialogue setting is both distinctive and normative, that is, participants are expected to confine themselves to the conversational actions allowed to their roles according to the dialogue genre (Heritage and Greatbatch, 1991, p. 106).

Further examples of departures from the norm are referred to as **violative talk**. Greatbatch (1986) studied agenda shifting procedures: a mechanism employed by interviewees to divert the topic of an interview towards their own agenda. Clayman (1988) analysed issues with interviewer neutrality, such as bias and subjectivity.

From a blend of social psychology and political science, following Harris (1991), Bull and colleagues (Bull and Mayer, 1993; Bull, 1994) studied instances of **equivocation**, systematically classifying the different ways in which politicians fail to answer questions properly. This work was later extended to include interviewers' face-threatening actions among the causes for equivocation (Elliott and Bull, 1996; Bull, 2003; Bull, 2008), identifying a class of questions called **avoid-avoid** for which any answers would result in a negative outcome for the interviewee, triggering the need of avoidance.

Finlayson (Finlayson, 2001) analyses the problem of the political interview as the clash between two worlds: the political world of discussion, negotiation and compromise *versus* the journalistic world of the forensic pursuit of truth. The clash Finlayson poses as a fundamental problem, is precisely why political interviews exist in the first place. This is also the reason why we have chosen this as the domain for our empirical investigations: a situation, conflictive in nature, to which two participants come voluntarily, agreeing on their goals in the large (spreading information) but potentially not in the small (what information to spread and how). This often forces them to consider carefully the course of the conversation and often leads to the deviations identified and analysed by researchers in the literature above, that is to what we here understand by instances of linguistic non-cooperation.

2.6 Summary

Cooperation in dialogue has been studied in different disciplines, from philosophy and the pragmatics of language, to conversation analysis, computational linguistics and dialogue systems design. Philosophers and theoretical linguists have claimed that participants in a conversation adopt a set of principles in order to maximise the efficiency and effectiveness of communication.

Computational linguists and dialogue systems designers have traditionally taken these assumptions as preconditions for the successful completion of an underlying task which conversation generally supports.

However, there is a recurrent distinction in the literature between cooperation at the level of the task supported by the conversation and cooperation at the level of the linguistic exchange itself. Although these kinds of cooperation are often related, in strict terms participants can cooperate or fail to cooperate independently at either level. This allows us to focus on cooperation at the level of the linguistic exchange, especially on the often neglected case of non-cooperation. Although some research has been done in task-level non-cooperation, which focuses on situations in which the task-related goals of the participants are at odds, it has been generally assumed that participants still cooperate at the linguistic level.

Efforts in computational linguistics and dialogue system design have tried to model non-cooperative behaviour scenarios, including negotiations and other instances of strategic talk, to a level of precision that gives rise to computational models. Unfortunately, the instances of non-cooperative behaviour in these approaches have been modelled explicitly, by means of specific rules for non-cooperation, which has limited the scope and breadth of phenomena they account for. Previous approaches to modelling dialogue without relying on full cooperation between the parties are based on rather complex descriptions of the interlocutors (user models), regard non-cooperative behaviour only as deception or account for exchanges that, yet adversarial, are cooperative in the sense that participants are following rules or strategies that describe their expected behaviour.

Still, research in these areas has lead to a number of concepts with a potential to address non-cooperative behaviour, namely dialogue games and discourse obligations. The key point is thus to abandon the direct rule-based approach for modelling non-cooperation explicitly and take a perspective in which expected, cooperative behaviour is specified by means of precise rules. Non-cooperation emerges then as a deviation from such expectations. This presupposes that there is a set of norms that define cooperative behaviour and that these norms can be explicated to the extent that they can be formalised and eventually implemented.

In order to do this, we look at the literature of political interviews and see that there is a set of underlying conventions that they try to adjust to. The empirical analysis of political dialogue suggests that such – usually implicit – norms are operative in conversation: if people deviate from these norms, their behaviour becomes an explicit topic of the conversation itself. The terminology used strongly suggests this, speaking of sanctionable behaviour, violative talk or equivocation to refer to these behaviours. Our focus is on non-cooperation in political interviews. We will use these insights to propose a set of explicit rules for political interviews, a global dialogue game, which then will be used in the rest of the thesis to systematically measure and model non-cooperation in terms of violations to the rules of the game.

Chapter 3

Non-Cooperation, Discourse Obligations and Dialogue Games

This chapter presents the conceptual framework used in the rest of the thesis for measuring and modelling non-cooperation in dialogue. It defines what we understand by non-cooperation in dialogue and demarcates the aspects of conversation we address in detail, discussing the simplifying assumptions for those we leave for future consideration. It also explains how discourse obligations and global dialogue games are combined to represent the norms and conventions associated with a specific conversational setting, ending with a detailed example from the domain of political interviews, which plays a central role in the rest of the thesis.

3.1 Overview of the Chapter

As discussed in Section 2.2, cooperation in dialogue can be observed at different levels of the interaction. In Section 3.2, we establish a clear distinction between cooperation in the activity of maintaining a conversation and cooperation in the performance of an underlying task. These are respectively referred to as **linguistic** and **non-linguistic cooperation**. The focus of this research is on linguistic non-cooperation and thereafter, unless explicitly stated, the terms "cooperation" and "non-cooperation" refer to the linguistic level.

A general **definition of linguistic non-cooperation** of a speaker with respect to a conversational setting is given in Section 3.3. We illustrate with examples from the domain of political interviews and motivate an exploration of the phenomenon in greater detail. We arrive at a compromise between an increased level of detail and the manageability of the research project by focusing on core speech acts. At the end of the section, we discuss briefly the assumptions in regard to other aspects of conversation that facilitated this compromise, such as turn-taking, grounding and non-verbal behaviour.

Section 3.4 posits **discourse obligations** and **global dialogue games** as suitable representations of the norms and conventions associated with a conversational setting. We discuss how these can be used to refine the definition of linguistic non-cooperation. A detailed dialogue game dealing with core speech acts in the political interview setting is described and formalised in Section 3.5. Later in the thesis, this device is used for measuring linguistic non-cooperation in naturally-occurring political interviews (Chapter 4) and for modelling conversational agents that can cope with and exhibit dialogue behaviour with different degrees of cooperation (Chapter 5).

3.2 Linguistic and Non-Linguistic Cooperation

As discussed in the previous chapter, participants in a dialogue can cooperate at different levels of the interaction. Here, we consider and illustrate the differences between two of these levels: cooperation in maintaining the flow of the conversation, for instance by asking appropriate questions at appropriate times and by providing relevant answers in response to those questions, and cooperation in achieving the goals of an underlying task, for instance by helping uncover the truth on a certain matter or by making important information available to a wider audience.

Dialogue generally supports an underlying activity. For the moment we focus on dialogue where this consists of a task or set of tasks that the participants aim to complete with the assistance of verbal interaction. Examples of such activities include the assessment of a student's knowledge via oral examination, the negotiation of the price of goods or services through bargaining, the gathering of evidence by means of courtroom cross-examination and many others.

The nature of the underlying activity and the associated social context impose constraints on the conversational behaviour considered acceptable for each dialogue participant. At the same time, conversational behaviour affects how successfully participants perform the underlying task. In the general case however, cooperation at one of these two levels of interaction does not directly translate to cooperation at the other level. This was illustrated earlier by the case of the witness under interrogation in a US trial, who could acceptably decline to answer a question by appealing to the Fifth Amendment thus being non-cooperative in relation with the goals of the cross-examination. Non-cooperation at the conversational level can result in lack of cooperation at the level of the task – take as an example the same witness remaining silent, rather than answering or appealing to the Fifth Amendment. However, there are cases in which unconventional or unexpected conversational behaviour can contribute in the completion of the underlying task. For instance, although sanctionable in courtroom cross-examinations, leading questions could help a sincere but forgetful witness remember important details that become evidence.

To illustrate the distinction further, recall the following fragment from Example 1.1 in Chapter 1:

Turn	**Speaker**	**Speech**
(1)	Paxman	(interrupting) Did you threaten to overrule him?
(2)	Howard	I, I, was not entitled to instruct Derek Lewis, and I did not instruct him.
(3)	Paxman	Did you threaten to overrule him?
(4)	Howard	The truth of the matter is that Mr. Marriott was not suspended. I-
(5)	Paxman	(overlapping) Did you threaten to overrule him?
(6)	Howard	-did not overrule Derek Lewis.

Howard's responses do not constitute relevant answers to the interviewer's question, nor do they work as explicit rejections. Therefore, they are non-cooperative at both levels of interaction, as they do not contribute to the common goals of the political interview task and also disrupt the dynamics of the conversation leading Paxman to pose essentially the same question over and over. Consider however, a fictional alternative in which Howard replies to the question by saying *'I will not answer that question, as it is not relevant to whether I exceeded the powers of my office'* would not be cooperative in terms of helping achieve the goals of an interview – widening the spread information, uncovering the truth, clarifying pressing issues, etc. However, it would be contributing at the linguistic level as it helps in preserving the flow of the conversation, for instance by triggering a sub-dialogue to solve the disagreement or by convincing the interviewer to drop the question and move on.

As detailed in the previous chapter, the distinction between linguistic and non-linguistic (also called task-related, high-level or social) cooperation has been addressed before. Attardo (1997) revisits Gricean pragmatics, relating non-linguistic cooperation to participants' behaviour towards realising task-related goals, and linguistic cooperation to assumptions on their respective behaviour in order to encode and decode intended meaning. From a computational perspective, Bunt (1994) relies on a similar distinction when defining dialogue acts. Also, Traum and Allen (1994) introduce discourse obligations as an alternative to joint intentions and shared plans, to allow for models of dialogues in which participants do not share the same high-level goals and where behaviour is also determined by “a sense of obligation to behave within limits set by the society” (Traum and Allen, 1994, p. 2). Walton and Krabbe (1995) proposed a typology of dialogue based on the initial situation triggering the exchange and participants' shared aims and individual goals. Based on their work, Reed and Long (1997) distinguish cases where participants follow a common set of dialogue rules and stay within a mutually acknowledged framework from a stronger notion in which their individual goals point in the same direction. Boella et al. (1999) discuss the role of intentions in cooperation and distinguish between dialogue and domain goals. Dialogue goals are further subdivided in conversational and linguistic goals, assuming that participants align their conversational goals avoiding to offend each other so that the dialogue proceeds smoothly. Linguistic goals relate to the production and interpretation of utterances.

With the above distinction clear, it must be noted that in the rest of the thesis we do not deal explicitly with non-linguistic cooperation. This means that we do not consider shared goals as part of our approach, nor do we include the reasoning of the participants in terms of each others' private goals. The task-related goals of each party are considered private and whether they point in the same direction or are in conflict is orthogonal to how we deal with linguistic non-cooperation. From this point on, for the sake of brevity and unless explicitly indicated, the terms **cooperation** and **non-cooperation** will refer to the linguistic level. Further, these notions apply also to cases in which the underlying activity cannot be described in terms of tasks and goals. Arguably, exchanges such as small talk and phatic conversation do not qualify as task-oriented dialogues. However, these types of dialogue do serve social purposes (e.g. for politeness or to increase familiarity between the parties) and therefore it is possible to identify a set of –usually tacit and culture-bound– rules or conventions on how to take part in them adequately allowing for the identification of non-cooperation regardless of the absence of tasks and goals.

3.3 Defining Non-Cooperation in Dialogue

Conversational behaviour can be accounted for either from a normative perspective in terms of how participants ought to behave, or from a descriptive perspective in terms of how participants do actually behave (Piwek, 2006). In a normative approach, the interaction is often characterised by means of rules that participants are expected to follow when they engage in conversation (Walton and Krabbe, 1995). These rules are determined by contextual aspects linked to, among others, the type of exchange, the purpose of the conversation and social conventions. In these terms, following the definitions in the previous section, participants are cooperative when they stick to the rules and stay within the restrictions imposed by these norms. Consistently, participants are non-cooperative when they break the rules or fail to observe the constraints associated with their respective roles in the type of exchange in which they engage.

Normative approaches lead to idealisations of the conversational settings they deal with. This is because larger sets of rules result in more complex models and the consequences of cumbersome models, such as interference among rules, their precedence and combined effects, are harder to track or predict. On the one hand, idealisations are desirable when the aim is mainly that of formalising phenomena for precise reasoning and automation. On the other hand, they inevitably lead to non-conforming interactions that fall outside of the model being left unaccounted for. In what follows we explain how, by relying on the idealisation of a dialogue setting, we account for interactions that deviate from the norm. That is, we describe how participants do actually behave by relying on a characterisation of how they ought to behave, thus combining in a same account elements from the descriptive and normative approaches to dialogue modelling.

Deviations, as discussed above, are what we call non-cooperation in dialogue. Here, we characterise a class of behaviours that occur in natural conversation and later in the chapter devise a set of rules which when broken by speakers result in non-cooperation. We therefore propose the following **operational definition**:

> The behaviour of a speaker is non-cooperative if it fails to comply with at least one of the rules that specify the speaker's role in the current conversational setting.

The instances of non-cooperative behaviour, that is whenever one of the rules that specify the speaker's role is broken, are called **non-cooperative features**. The number of occurrences of these features will then determine the extent of non-cooperation of each speaker in a given exchange. In this way, **non-cooperation in dialogue** is a matter of **degree**, ranging from total non-cooperation in which each of the speaker's actions fails to comply with at least one rule and total cooperation in which all their actions are in accordance with the rules specifying their role.

3.3.1 Non-Cooperation in Political Interviews

We illustrate the definition above by giving an example from the political interview conversational setting. In Section 2.5 of the previous chapter we discussed a range of empirical studies on the nature of political interviews. Heritage, Greatbatch and Clayman analyse the distinctive organization of news and political interviews and describe how the institutions associated to each speaker restrict their roles (Clayman, 1988; Heritage and Greatbatch, 1991; Heritage, 1998; Greatbatch, 1988; Clayman and Heritage, 2002; Heritage, 2005). They note that in practice these restrictions are occasionally not observed, but only as departures from the expected behaviour and often resulting in problematic and even sanctionable courses of action. These sanctionable courses of action correspond with what we defined above as non-cooperative behaviour.

These insights and the literature on the analysis of political interviews discussed in Chapter 2 indicate that there is a consensus regarding the speakers roles in terms of what is acceptable during the conversation and what is not. An operational definition of non-cooperative behaviour specific to the roles of interviewer and interviewee in political interviews can thus be as follows:

- **Non-cooperation of interviewers** with respect to the political interview conversational setting happens whenever they:

    1. express personal opinions;

2. argue, debate with or criticise the positions of the interviewee;
3. agree with, support or defend the positions of the interviewee;
4. criticise or attack the interviewee or their organisation;
5. after a response from the interviewee, they neither:
   - move on to the next question; nor
   - state that the response is not a relevant answer;
6. interrupt or speak simultaneously with the interviewee;
7. end the exchange abruptly or threaten to do so.

- **Non-cooperation of interviewees** with respect to the political interview conversational setting happens whenever they:

8. ask questions of the interviewer (with the exception of CRs);
9. provide incomplete answers;
10. make unsolicited comments;
11. initiate changes of topic;
12. criticise or attack the interviewer or their organisation;
13. after a question from the interviewer, they neither:
    - provide a suitable answer; nor
    - state that the question will not be answered;
14. interrupt or speak simultaneously with the interviewer;
15. end the exchange abruptly or threaten to do so.

For illustration, consider the fragments in Figures 3.1 and 3.2. The utterances of each speaker are numbered and annotated with the behaviours they exhibit from the list above.

Turn	Speaker	Utt.	Speech	Features
(11)	PAXMAN	IR.1	(overlapping) Did you threaten to overrule him?	6
(12)	HOWARD	IE.1	Mr. Marriot was not suspended.	10,13
(13)	PAXMAN	IR.2	Did you threaten to overrule him?	5
(14)	HOWARD	IE.2	I have accounted for my decision to dismiss Derek Lewis-	
(15)	PAXMAN	IR.3	(overlapping) Did you threaten to overrule him?	5,6
(16)	HOWARD	IE.2	(overlapping) -in great detail before the House of Commons-	10,13,14
(17)	PAXMAN	IR.4	I note that you're not answering the question whether you threatened to overrule him.	
(18)	HOWARD	IE.3	Well, the important aspect of this which it's very clear to bear in mind-	11,13
(19)	PAXMAN	IR.5	(interrupting) I'm sorry, I'm going to be frightfully rude but-	6
(20)	HOWARD	IE.4	(interrupting) Yes, you can-	14
(21)	PAXMAN	IR.6	(overlapping) I'm sorry-	6
(22)	HOWARD	IE.4	(overlapping) - you can put the question and I will give you, I will give you an answer.	
(23)	PAXMAN	IR.6	(overlapping) -it's a straight yes-or-no question and a straight yes-or-no answer:	
		IR.7	did you threaten to overrule him?	
(24)	HOWARD	IE.5	I discussed the matter with Derek Lewis.	
		IE.6	I gave him the benefit of my opinion.	
		IE.7	I gave him the benefit of my opinion in strong language, but I did not instruct him because I was not, er, entitled to instruct him.	
		IE.8	I was entitled to express my opinion and that is what I did.	9
(25)	PAXMAN	IR.8	With respect, that is not answering the question of whether you threatened to overrule him.	
(26)	HOWARD	IE.9	It's dealing with the relevant point which was what I was entitled to do and what I was not entitled to do,	11,13
		IE.10	and I have dealt with this in detail before the House of Commons and before the select committee.	10

Figure 3.1: Fragment of the Paxman-Howard interview (Example 1.1 in Chapter 1) annotated with the behaviours in the definition of non-cooperation in political interviews. The numbers in the right-most column correspond with those in the definition on Page 64.

Turn	Speaker	Utt.	Speech	Features
(01)	PAXMAN	IR.1	We're joined now from his count in Bethnal Green and Bow by George Galloway.	
		IR.2	Mr Galloway, are you proud of having got rid of one of the very few black women in Parliament?	4
(02)	GALLOWAY	IE.1	What a preposterous question.	
		IE.2	I know it's very late in the night, but wouldn't you be better starting by congratulating me for one of the most sensational election results in modern history?	8,11
(03)	PAXMAN	IR.3	Are you proud of having got rid of one of the very few black women in Parliament?	4,5
(04)	GALLOWAY	IE.3	I'm not- Err, Jeremy, move on to your next question.	
(05)	PAXMAN	IR.4	You're not answering that one?	
(06)	GALLOWAY	IE.4	No, because I don't believe that people get elected because of the colour of their skin. I believe people get elected because of their record and because of their policies. So move on to your next question.	
(07)	PAXMAN	IR.5	Are you proud-	4,5
(08)	GALLOWAY	IE.5	(Interrupting) Because I've got a lot of people who want to speak to me.	10,14
		IE.6	If you ask that question again, I'm going, I warn you now.	15
(09)	PAXMAN	IR.6	Don't try and threaten me Mr Galloway, please.	5
			(. . .)	
(28)	GALLOWAY	IE.7	You are actually conducting one of the most, even by your standards, one of the most absurd interviews I have ever participated in.	12
		IE.8	I have just won an election.	10
		IE.9	Can you find it within yourself to recognise that fact?	8, 11
		IE.10	To recognise the fact that the people of Bethnal Green and Bow chose me this evening.	10,11
		IE.11	Why are you insulting them?	8
(29)	PAXMAN	IR.7	I'm not insulting them, I'm not insulting you	1,5
(30)	GALLOWAY:	IE.12	You are insulting them, they chose me just a few minutes ago.	
		IE.13	Can't you find it within yourself even to congratulate me on this victory?	11
(31)	PAXMAN:	IR.8	Congratulations, Mr Galloway.	5
(32)	GALLOWAY:	IE.14	Thank you very much indeed. *(Waves, removes microphone and leaves)*	15

Figure 3.2: Fragment of the Paxman-Galloway interview (Example 1.3 in Chapter 1) annotated with the behaviours in the definition of non-cooperation in political interviews.

Preliminary studies based on a similar definition lead to encouraging results (Plüss, 2009; Plüss, 2010). A fully cooperative dialogue is regarded as the ideal situation in which participants always do what is best to preserve the functioning of the conversation. In a political interview, this would refer to cases in which questions are answered directly and in a complete way, these answers are accepted without comments by the interviewer who goes on with the next question on the agenda, interlocutors speak in turn without interruption or overlapping, etc. The degree of non-cooperation of each participant was measured as the ratio between the utterances exhibiting one or more of the behaviours in the definition and the total number of utterances of the speaker. The rationale behind this measure was to capture the proportion of cases in which a speaker had made a decision on what to say and the contribution displayed at least one of the behaviours defined as non-cooperative. In the example in Figure 3.1, for instance, these values are $5/8 = 0.625$ for the interviewer (Paxman) and $7/10 = 0.7$ for the interviewee (Howard). The respective values for the interviewer and interviewee in the fragment in Figure 3.2 are $6/8 = 0.75$ and $10/14 = 0.714$. Subsequent attempts to apply the scheme reliably on a larger corpus of interviews, however, showed that the definition required a higher level of detail. Some of the shortcomings include:

- Ambiguity: for example, a personal opinion expressed by the interviewer in support of the views of the interviewee could be classed as 1 or 3 (or both);
- Non-exclusive categories: an interviewee avoiding a question by shifting the agenda, for instance, falls under three categories (10, 11 and 13), and even four if the shift is introduced by asking a question of the interviewer (8);

- Imprecision: for example, it is not clear what constitutes a question or an answer, the criteria for determining relevance and suitability of the contribution are underspecified, etc.;

- False positives: interruptions can be cooperative, for instance, when the listener detects an error or misunderstanding in the speaker utterance, a timely interruption can help preserve the flow of the conversation[1];

- Forced ratio: utterances are not the best unit for computing the degree of non-cooperation as one utterance can exhibit more than one of the behaviours in the definition;

- Coarse granularity: all the behaviours contribute the same towards non-cooperation, with an interruption, for instance, being equally non-cooperative as an attack or the covert avoidance of a question.

3.3.2 Towards a Revised Definition: Focusing on Core Speech Acts

The definition above covers several aspects of the actions participants make in conversation: from the organisation of turns, to the establishment of a common ground and the type of function and propositional content associated with contributions. Put in terms of dialogue acts, according to the classification used by Poesio and Traum (1997; 1998) with elements from Conversation Acts Theory (Traum and Hinkelman, 1992) and the Discourse Resource Initiative (Initiative, 1997; Allen and Core, 1997) the rules apply at one or more of the following levels:

[1]For other cases of cooperative interruptions, see for instance the work of Aist (1998) and Mostow et al. (2003) on an automated reading tutor, and of Smith et al. (2011) on companion ECAs.

Turn-Taking: regarding interruptions, overlapped speech and closings;

Grounding: indicating what constitutes evidence of understanding, such as adequate continuations;

Core Speech Acts: restricting the functions (questions, answers) and contents (adequate, relevant) that are expected from the role of each participant;

Argumentation Acts: dealing with the global flow of the conversation and the sequences that are considered adequate for the type of interaction.

In order to deepen the level of detail in the definition of non-cooperation, we will not be dealing with aspects of turn-taking and grounding and instead focus on core speech and argumentation acts. This means that we make the following **simplifying assumptions**[2]:

- participants do not interrupt each other nor do they speak simultaneously;
- non-verbal behaviour and prosodic features do not affect the function and content of core speech acts;
- grounding is always achieved correctly, that is participants understand flawlessly at all times what the other party contributed and provide adequate evidence indicating that this is the case;
- the global argumentation act is always one and the same: the political interview, we will not consider subdialogues or transitions into dialogues of a different type (e.g. clarification, negotiation, confrontation, debate).

[2] These phenomena are ignored when analysing naturally-occurring dialogue.

These assumptions set the scope for the results in the thesis and open directions for future research. From the point of view of a speaker holding the floor, interruptions and overlapped speech require monitoring, reasoning and responding to the behaviour of the other party even when they do not hold the floor. For instance, a speaker who decides to stop talking when interrupted by the other interlocutor would be more cooperative than one deciding to continue speaking, raising the voice, etc. The generation or prediction of interruptions requires that the listener interprets the utterances of the speaker while they are being produced so as to decide on whether to start a contribution before receiving the floor. Current work along these directions includes the treatment of barge-in interruptions for companion embodied conversational agents (Crook et al., 2010; Smith et al., 2011) and the flourishing area of research know as incremental dialogue processing (DeVault et al., 2009; DeVault et al., 2011; Schlangen and Skantze, 2011; Rieser and Lemon, 2011; Traum et al., 2012).

The processes involved in establishing a common ground (Clark and Brennan, 1991; Traum, 1994) lend themselves to displays of non-cooperation. In the presentation phase (Clark and Brennan, 1991), for instance the use of ambiguous utterances and implicatures could lead to defensible or non-committal contributions (Benotti, 2009; Benotti, 2010; Asher, 2012; Asher and Lascarides, in press) and become sources of mismatches in the common ground. In the acceptance phase, the evidence of understanding could also be misleading or ambiguous. For example, an interviewee providing ambiguous indication of their understanding of a question but responding with an irrelevant comment could defensibly claim that the question had been misunderstood if the response is challenged by the interviewer. Recent research on degrees of grounding (Roque and Traum, 2008; Brown-Schmidt, 2012) and on listener feedback (Buschmeier and Kopp, 2012; Bunt, 2012) are of relevance for future work on non-cooperation along these lines.

Clarification subdialogues (Schegloff, 1987; Purver, 2004) are related to grounding and occur in most types of interaction. These exchanges are governed by their own set of rules, often independently from those in the containing dialogue. They should be considered separately from the normal flow of the conversation and should be analysed based on their own rules. For instance, an interviewee asking a question of the interviewer as a request for clarification would not be considered in violation of the rules of political interviews. As we assume grounding is flawless, we will not deal with cooperation in clarification subdialogues here. When present in naturally occurring conversation, as in Turns (05) and (06) in Figure 3.2, these exchanges are not annotated. Recent research on the analysis and generation of clarification requests (Rieser and Moore, 2005; Purver, 2006; Schlangen and Fernández, 2007; Skantze, 2007; Benotti, 2009; Ginzburg, 2012; Cooper, 2013) will be relevant to future work in order to account for these subdialogues.

Finally, to extend the approach to other types of dialogue (Walton and Krabbe, 1995; Traum et al., 2008b) would involve devising adequate explicit or implicit rules and conventions that govern the exchange. Recent and current research in computational linguistics and dialogue systems has focused on a great many such types: negotiation (McBurney et al., 2003; Traum et al., 2005; Traum, 2008; Plüss et al., 2011), argumentation and persuasion (Piwek, 2008; Georgila et al., 2011), tutoring (Lu et al., 2007; Litman et al., 2010; Dzikovska et al., 2010; Dzikovska et al., 2012), expository dialogues (Stoyanchev and Piwek, 2010b; Stoyanchev and Piwek, 2010c), courtroom interrogation (Asher et al., 2012; Asher and Lascarides, in press), etc.

We will return to the issues above when describing future lines of research in Chapter 6. Core speech acts were chosen to analyse the definition in greater depth as they are of salient relevance to dialogue management, which is the focus of Chapter 5 when modelling non-cooperation in conversational agents. Core speech acts also allow for a rich set of well-defined phenomena specific to political interviews to be analysed in the context of non-cooperation. These are the focus of the remainder of this chapter and of the empirical studies described in Chapter 4.

3.4 Discourse Obligations and Dialogue Games

In this section, we propose a combination of discourse obligations (Traum, 1994) and global dialogue games (Hamblin, 1970; Walton and Krabbe, 1995) as a device to represent the rules and conventions associated with a dialogue type. A global **dialogue game** is a description of the conversational situation. For each speaker role, it specifies what actions are allowed (and therefore which ones are not), what contents are adequate for each contribution, what actions the participants are conventionally obliged to perform as a consequence of actions that have been performed earlier in the dialogue by themselves or by the other party, and so forth.

Contrary to other uses of the term (Levin and Moore, 1977; Carlson, 1983; Mann, 1988; Carletta et al., 1997; Pulman, 2002), dialogue games here refer to the entire dialogue and not just to local, goal-directed exchanges, such as a question-answer pair[3]. Hamblin (1970) studied fallacies using similar descriptions, which he called **dialectical systems**. Fallacies result from speakers breaking the rules of the dialectical system for valid ar-

[3]These local **dialogue games** and closely related notions have also been called **conversational procedures** (Power, 1979), **conversational games** (Kowtko et al., 1992; Pulman, 1997; Lewin, 2000) and **dialogue macrogames** (Mann, 2002).

guments. The approach was extended to the analysis of informal logic and argumentation dialogues by Walton and Krabbe (1995). They proposed a classification of dialogue types based on the initial situation, the individual and shared goals of the participants and mechanisms by which they can switch licitly between dialogues of different types[4].

Along these lines, a global dialogue game can be thought of as a contract participants subscribe to by agreeing on a specific type of interaction. At the core of the game there is a set of rules that capture the conventions associated with the dialogue type. Participants are expected to act according to these rules. Traum and Allen (1994) first pointed at this "social pressure" and proposed **discourse obligations** as an alternative to joint intentions (Cohen and Levesque, 1991), shared plans (Grosz and Sidner, 1990) and conversational games (Power, 1979; Kowtko et al., 1992) to allow for models of dialogue in which participants do not have the same task-level goals and therefore would not always adopt each other's intentions, agree on a shared plan or accept to engage in a particular (local) conversational game. In this view, behaviour is not only determined by participants' goals, but also by "a sense of obligation to behave within limits set by the society that the agent is part of" (Traum and Allen, 1994, p. 2). The typical example is that of someone responding at all to a question they do not wish to answer instead of remaining silent.

Discourse obligations have been used for modelling several aspects of dialogue management (Jameson and Weis, 1995; Matheson et al., 2000; Kreutel

[4]Walton and Krabbe (1995) discussed in detail the rules for argumentation dialogues. Several authors have based on their approach to model other types of conversation, including human-computer debate (Maudet and Moore, 2001), computer-mediated crosslingual communication (Piwek et al., 2007) and purchase negotiation, argumentation and conflict resolution in agent communication for multi-agent systems (McBurney and Parsons, 2009; Karunatillake et al., 2009)

and Matheson, 2003b)[5]. However, as far as we know, discourse obligations have always been used in connection with general phenomena of conversation. We take this one step further and think of obligations as determined by the type of dialogue, rather than fixed for all interactions. That is, the correct handling of obligations is determined by the dialogue game.

Discourse obligations then follow naturally from the rules of global dialogue games and are associated with linguistic cooperation in the following way: when obligations are honoured, the rules of the game are observed and the result is linguistic cooperative behaviour. Non-cooperative behaviour is thus by definition beyond the descriptive reach of dialogue games. Adding explicit rules to account for the variations present in non-cooperative conversations would require an additional set of rules for each possible unconventional behaviour. In this framework, we focus on cases in which obligations are not honoured, much in the same way as Hamblin focused on breaking the rules of valid dialectical arguments in the analysis of fallacies, rather than considering explicitly all the possible ways in which a fallacy could take place (Hamblin, 1970, Chapter 1).

3.4.1 Formalisation

Below we formalise the elements that make up a dialogue game.

Action Labels

The rules of a dialogue game are formulated in terms of the actions that participants perform during a conversation. These actions will be represented as labels that capture those aspects of the speaker's contributions that

[5]Kreutel and Matheson (2000) showed how from a model of dialogue based on obligation it is possible to reconstruct the structures used by intention-based approaches. In the same paper they hint at the possibility of using obligations – instead of, e.g. intentions and politeness (Boella et al., 1999) – to study cooperation in dialogue.

are necessary for applying the rules. They result from an analysis of the function and content of the contributions, the details of which are presented and discussed at length in Chapter 4.

For the time being, we assume that dialogue transcripts have been annotated according to a taxonomy T of conversational actions, so that each turn is associated with a sequence of labels that characterise the speaker's actions in the turn. Chapter 4 presents a coding scheme for naturally-occurring political interviews and shows how annotations in that coding scheme can be mapped to the action labels used here. The labels that represent conversational actions have the following structure:

$$(\langle \mathit{ID} \rangle) : \langle \mathit{action\ name} \rangle [@(\langle \mathit{referent\ ID} \rangle)]\ \ [\langle \mathit{binary\ flags} \rangle]$$

In their simplest form, action labels are formed by an identifier in parenthesis, $\langle \mathit{ID} \rangle$, and the name of the action, $\langle \mathit{action\ name} \rangle$. The elements in square brackets are optional:

- $\langle \mathit{referent\ ID} \rangle$ is the identifier of a previous action referred to by the current action, and
- $\langle \mathit{binary\ flags} \rangle$ are action-specific binary flags, examples include whether the action is repeated or new in the dialogue, whether a reply to a question is complete or incomplete, etc.

Labelled Dialogue

A labelled dialogue D is a sequence of turns $\langle t_1; \ldots; t_n \rangle$. Each turn t_i is a pair (s_i, L_i), where s_i is the speaker and L_i is a sequence of instances of action labels with adequate values for the parameters (identifier, references and binary flags as applicable). Speakers alternate in each turn, so $s_i \neq s_{i+1}$, for $i \in \{1 \ldots n-1\}$. Also, further restrictions on what labels can appear associated with each speaker in a correctly annotated dialogue can be specified for pairs (s_i, L_i).

Obligations

As mentioned above, obligations are actions that participants are expected to perform as soon after their introduction as possible. Assuming S is the set of speakers, we represent obligations as pairs (s, l) in which $s \in S$ is the speaker upon whom the obligation is imposed and l is an action label.

Dialogue Game

A dialogue game G is a triple $(\mathit{Allow}, \mathit{Introduce}, \mathit{Discharge})$, where each element is a set of rules, respectively, of the form:

(a) $[s : L]$, where s is a speaker and each L is a subset of the labels in the annotation taxonomy. Such a rule specifies that speaker s is obliged to limit his or her actions to those in L.

(b) $[(s, l) \rightsquigarrow o]$, where s is a speaker, l is a label in the taxonomy and o is an obligation. Such a rule specifies that action l, if performed by speaker s in the dialogue, creates an obligation $o = (s_o, l_o)$ that specifies that speaker s_o is expected to perform action l_o.

(c) $[l_a \succ l_b]$, where l_a and l_b are labels. Such a rule specifies that action l_a implicitly performs action l_b. These rules define a binary relation $\succ: T \times T$ that we assume reflexive, so for any $l \in T$, $l \succ l$. Together, these rules define the way obligations are discharged.

Discourse obligations are mandates on the speakers that constrain and motivate conversational behaviour. They are updated, that is introduced and discharged, dynamically along the course of the conversation. Obligations are to be discharged as soon as possible. This means that if a speaker's actions impose any obligations on him or herself, these are expected to be discharged within the same turn. On the other hand, if the speaker's actions impose obligations on the other speaker, these should be discharged in the turn immediately after.

We can distinguish between **static obligations** and **dynamic obligations**. Static obligations are those that limit what actions participants are allowed to take according to their role in the conversation as specified by the rules in *Allow*. Dynamic obligations are introduced and discharged along the dialogue as a consequence of the actions that participants actually take. They are introduced according to the rules in *Introduce*, and discharged according to those in *Discharge*.

3.4.2 A Revised Definition of Non-Cooperation in Dialogue

In the context of the definitions above, linguistic non-cooperation of a dialogue participant with respect to a conversational setting equates to the participant violating the rules of the dialogue game for that conversational setting. From this perspective, each turn in a dialogue is associated with an amount of cooperation and an amount of non-cooperation. The amount of cooperation is given by the number of dialogue rules directly conformed with in the turn, whereas non-cooperation is determined by the number of rules directly violated in the turn. We refer to instances in which rules are conformed with as **cooperative features** and to instances in which rules are broken as **non-cooperative features**. These notions, extended to all the turns of a participant, lead to the amount of cooperation and non-cooperation in the entire dialogue.

Participants can break the rules of the game in two ways: (a) by performing a conversational action that is not allowed for their role and (b) by failing to perform an action they were obliged to perform. Instances of (a) are violations of static obligations, which we call **static non-cooperative features**. Instances of (b) are violations of dynamic obligations, which we call **dynamic non-cooperative features**. An analogous distinction is made for cooperative features, called, respectively, **static cooperative features** and **dynamic cooperative features**.

The **degree of cooperation** of each dialogue participant is thus the ratio between the number of cooperative features – static and dynamic – and the total number of features of that participant. In general, this value can be obtained for the entire conversation and for any continuous fragments.

In the rest of the section, we formalise the concepts above. Chapter 4 describes a method to compute the dynamic obligations of participants in each turn of a dialogue. It also explains how to compute static and dynamic cooperative and non-cooperative features and the degree of non-cooperation.

Dynamic Obligations

Formally, for a dialogue $D = \langle t_1; \ldots; t_n \rangle$, where each t_i is a turn, we represent dynamic obligations as a sequence $O_D = \langle O_{D,0}; O_{D,1}; \ldots; O_{D,n} \rangle$, where each $O_{D,i}$ is a list $\langle o_1, \ldots, o_k \rangle$ of the obligations after turn t_i, and $O_{D,0}$ is the list of obligations before the dialogue starts. As before, each obligation o_i is a pair (s_i, l_i), where s_i is a speaker and l_i is an action label.

Obligations are **updated** in each turn of the dialogue. This means that, for each turn t_i, $O_{D,i}$ is computed based on $O_{D,(i-1)}$ by **discharging** existing obligations and **introducing** new ones[6]:

[6]Recall that the implicit performance relation $\succ$ is reflexive, so for any label l it is $l \succ l$.

- An obligation $o = (s, l) \in O_{D,i-1}$ is **discharged** in turn $t_i = (s_i, L_i)$ if $s = s_i$ and $l_j \succ l$ for some label $l_j \in L_i$. To discharge several obligations in the same turn, we generalise this definition in the obvious way.

- An obligation o is **introduced** in turn $t_i = (s_i, L_i)$ if there is a rule $(s, l) \rightsquigarrow o$ in the game such that $s = s_i$ and $l_j \succ l$, for some label $l_j \in L_i$ (meaning that obligations are introduced by implicitly performed actions). To introduce several obligations in the same turn, we generalise this definition in the obvious way.

Those obligations in $O_{D,(i-1)}$ that are not discharged in turn t_i are carried over to $O_{D,(i)}$.

Cooperative and Non-Cooperative Features

As said above, cooperative and non-cooperative features are instances of, respectively, observed and neglected static and dynamic obligations:

Static cooperative features: actions performed by the speaker that are allowed for his or her role in the dialogue.

Static non-cooperative features: actions performed by the speaker that are disallowed for his or her role in the dialogue.

Dynamic cooperative features: obligations on the speaker that were discharged in the current turn.

Dynamic non-cooperative features: obligations on the speaker that were not discharged in the current turn.

Formally, for a dialogue $D = \langle t_1; \ldots; t_n \rangle$ features are grouped in two sequences, $SF_D = \langle sf_1; \ldots; sf_n \rangle$ and $DF_D = \langle df_1; \ldots; df_n \rangle$, of static and dynamic features, respectively. The elements in both sequences are triples (s_i, C_i, NC_i) where s_i is the speaker in turn t_i, C_i is the list of cooperative features and NC_i is the list of non-cooperative features (static for sf_i and dynamic for df_i).

Static Features. These are determined by checking, in each turn, whether the actions performed by the speaker are allowed for his or her role as specified in the dialogue game. If an action is in the the speaker's set of allowed actions, then it constitutes a static cooperative feature, otherwise it becomes a static non-cooperative feature. Formally, this means that in turn $t_i = (s_i, L_i)$, for each $l \in L_i$ if $l \in L$, with $[s_i : L]$ a dialogue game rule in $G(1) = \mathit{Allow}$, then l is an element of C_i in $sf_i = (s_i, C_i, NC_i)$. Otherwise, l is an element of NC_i.

Dynamic Features. These are determined by the speaker's actions and obligations in each turn. If an obligation on the speaker has been discharged within the turn, then it constitutes a dynamic cooperative feature, otherwise it becomes a dynamic non-cooperative feature. Formally, this means that for turn $t_i = (s_i, L_i)$, an obligation $o = (s_o, l_o) \in \big(O_{D,(i-1)} \setminus O_{D,i}\big)$, for which $s_o = s_i$ has been discharged in the current turn, so l_o is a dynamic cooperative feature and thus an element of C_i in $df_i = (s_i, C_i, NC_i)$[7]. On the other hand, an obligation $o = (s_o, l_o) \in O_{D,i}$ for which $s_o = s_i$ has not been discharged in the current turn, so l_o is a dynamic non-cooperative feature and thus belonging in NC_i.

[7]Obligations that are introduced and discharged in the same turn are an exception to this rule, as they will not show up in $O_{D,(i-1)}$ nor in $O_{D,i}$, but they still constitute dynamic cooperative features. In the implementation of the method described in Chapter 4 we address this by using an additional structure to handle such cases.

Degree of Non-Cooperation

With the static and dynamic features for each turn, we can take the proportion of them that are non-cooperative as an indicator of the extent to which each participant acted outside the rules of the game. This is the **degree of non-cooperation** of a dialogue participant with respect to a dialogue game. Formally, the degree of non-cooperation of speaker s in dialogue D is:

$$dnc_{D,s} = \frac{ncf_{D,s}}{cf_{D,s} + ncf_{D,s}}$$

where $cf_{D,s}$ is the number of cooperative features – both static and dynamic – of participant s and $ncf_{D,s}$ is the analogous for non-cooperative features. This is[8]:

$$cf_{D,s} = \sum_{\substack{i=1 \\ [s_i=s]}}^{n} |sf_i(2)| + |df_i(2)|$$

$$ncf_{D,s} = \sum_{\substack{i=1 \\ [s_i=s]}}^{n} |sf_i(3)| + |df_i(3)|$$

Note that, although these definitions are here expressed for the entire dialogue, the same applies if any subsequence of turns is considered.

3.5 A Dialogue Game for Political Interviews

This section presents a dialogue game for the core speech acts in political interviews. We start by specifying the conversational actions that dialogue participants do during the conversation by giving the set of possible action

[8]Recall that the elements in the sequences of static and dynamic features $SF_D = \langle sf_1; \ldots; sf_n \rangle$ and $DF_D = \langle df_1; \ldots; df_n \rangle$ are triples (s_i, C_i, NC_i), where s_i is the speaker in turn t_i, and C_i and NC_i are the associated lists of, respectively, cooperative and non-cooperative features. Components of triples are accessed by indexation, so for instance $|sf_i(2)|$ denotes the length of the second component of triple sf_i, i.e. of the list of static cooperative features C_i of the i-th turn.

labels. Later, we present the rules for the political interview dialogue game, grounding them on the literature and providing examples. We formalise these concepts and ultimately present a dialogue game structure used in the rest of the thesis to illustrate the approach to measuring and modelling linguistic non-cooperation in dialogue applied to political interviews.

3.5.1 Action Labels for Political Interviews

The labels in political interviews have the following structure:

$(\langle ID\rangle)$: $[\langle validity\ prefix\rangle$-$]\langle action\ name\rangle[@(\langle referent\ ID\rangle)]$ $[\langle repeated\ \&\ complete\ flags\rangle]$

Note that the differences with respect to general definition presented in the previous section are an optional validity prefix as part the action name and two specific binary flags. Below we describe these elements in further detail and provide examples.

Identifiers and Action Names

Instances of action labels are identified by numbering them sequentially as they appear in the dialogue, so $\langle ID\rangle \in \mathbb{N}$.

Action names can take one of five possible values:

- `statement`: an assertion that conveys information and is not in response to a request from the other party. Examples include preambles to questions and voluntary comments unrelated to the topic of the interview or to the question under consideration.

- `question`: a request for the other party to provide some information. Examples include direct questions, and assertions that invite comments or opinions[9].

[9] As stated in Section 3.3.2, clarification requests are not considered questions in this study and are left unlabelled.

- `reply`: an assertion providing information in response to a request from the other party. Examples include complete and partial replies to questions, as well as assertions that although made in response to a question fail to provide any of the information requested without explicitly declining to answer[10].

- `acceptance`: a positive reaction to a previous contribution from the other party. This means explicitly agreeing with an assertion or to a request or contribution the other participant has made. Examples include agreeing to answer a question, or stating that a reply successfully answers a question.

- `rejection`: a negative reaction to a previous contribution from the other party. This means explicitly contradicting an assertion, refusing to agree to a request or dismissing an answer as inadequate. Examples include, declining to answer a question, stating that a reply is not informative enough, etc.

Valid and Invalid Actions

Some of the actions are further classified as valid or invalid depending on their content, on the context in which they take place and on the role of the speaker. A qualifier prefix is added to the action name indicating whether the action is valid or invalid. We distinguish valid from invalid actions for all statements, questions and replies, but not for acceptances and rejections. The criteria for validity are based on the discussion on the analysis of political interviews in Chapter 2.

[10] Responses to clarifications requests (e.g. verbatim repetition of previous contributions or rephrases) are not considered replies in this study and are left unlabeled.

Statements. These informative actions are often used by interviewers to set up the context for a question, for instance by providing facts and figures or quotations. The criteria for validity are as follows:

- `valid-statement`: new assertions by the interviewer, in which the information provided is objective and relevant to the topic of the interview, and in which any quotations and figures are accurate. In the (rare) cases of interviews that start with a statement from the interviewee, these too are considered valid, provided that the information conveyed is accurate and relevant to the topic of the interview[11].

- `invalid-statement`: any assertions by the interviewee that are not in response to a question by the interviewee, with the exception of those in the first turn of interviews that start with a statement by the interviewee, which are invalid only if they are either repeated, inaccurate or irrelevant to the topic of the interview. Also, any assertions made by the interviewer that are either repeated, subjective, inaccurate or irrelevant to the topic of the interview or of the line of questioning.

Questions. These requests for information are used by interviewers to elicit information from the interviewee and to set the topical agenda. The criteria for validity are as follows:

- `valid-question`: requests for information made by the interviewer, which are neutral, relevant to the current state of affairs and which the interviewee is in a position to answer.

- `invalid-question`: any questions posed by the interviewee[12]. Also,

[11]Note that objectivity is not a requirement for the validity of statements made by the interviewee.

[12]With the exception of clarification requests that are left unlabelled.

information requests by the interviewer, which have no answers or which the interviewee is not in a position to answer, requests for personal, private or irrelevant information and loaded questions which convey controversial assumptions, accusations or personal attacks.

Replies. These assertions are used by interviewees to provide the information required by interviewers. The criteria for validity are as follows:

- `valid-reply`: any new assertions made by the interviewee in response to a question from the interviewer, which provide (all or part of) the information requested.
- `invalid-reply`: any replies by the interviewer in response to questions from the interviewee[13]. Also, assertions made by the interviewee in response to a request for information, which are irrelevant, inaccurate or repeated.

As for the rest of the elements in the dialogue game, the distinction between valid and invalid contributions is grounded on social conventions and current practices in political interviews. It is worth noting, however, that validity is functional to classifying the contents of the actions of the dialogue participants, much in the same way as we distinguish between statements and questions based on an action's function. Validity therefore does not have a direct connection with cooperation as the latter depends on the obligations resulting from rules of the dialogue game. It is possible to think of conversational situations in which invalid actions are among the licit options allowed to the speakers – and sometimes even expected from them[14].

[13] Again, except for responses to clarification requests, which are left unlabelled.

[14] As an extreme example, BBC Radio 4 show *The Unbelievable Truth* is described by the broadcaster as "the game show in which panellists are encouraged to tell lies"

As mentioned in Chapter 2, at the end of Section 2.5, in hybrid political interviews (Hutchby, 2011) the neutrality requirement for interviewers is dropped, which would allow for (at least some) invalid questions to be licit in discharging obligations and therefore instances of cooperative behaviour.

References

The labels of responsive actions, that is replies, acceptances and rejections, include a reference to the action label they respond to. A reply action label, for instance, would include a reference to the question it is an answer to. As references point to previously identified actions, they too are represented as natural numbers. A restriction applies, as referents must appear earlier in the dialogue, so $\langle \textit{referent ID} \rangle \in \{1, \ldots, \langle \textit{ID} \rangle - 1\} \subset \mathbb{N}$, where $\langle \textit{ID} \rangle$ is the identifier of the current action. Furthermore, we only consider responses to actions of the other speaker, so references are limited to the identifiers of the action labels performed previously by the other party.

Binary Flags: Repetition and Completeness

Two aspects of the contents of conversational actions escape the classifications of validity discussed above and are treated at a broader level by the dialogue game: whether a question is repeated and whether a reply to a question is complete, in the sense that it provides all the information requested.

New and Repeated Contributions. The role of **repetition** in dialogue is an area of research in itself. Bazzanella (2011) notes that a characteristic of repetition, from a pragmatic perspective, is its polyvalence. Repeating

(`http://www.bbc.co.uk/programmes/b007mf4f`, last accessed: May 2013). Players take turns to lie about a given subject, trying to include fragments of truth without being detected by the other panellists.

utterances can serve different purposes and have diametrically opposite effects (e.g. express agreement or disagreement). In political interviews, we pay attention to self-repetition – speakers repeating something they themselves said earlier in the dialogue – that is intentional and that is not in response to a clarification request. Bazzanella refers to this as **intended redundancy** and identifies functions such as stressing a point, increasing emphasis and marking agreement or disagreement. We limit our study to intended self-repetition of questions and replies.

For replies, repetition is one of the causes of invalidity as indicated above. This is because, in ideal circumstances – what the dialogue game ultimately describes – once relevant and sufficient information has been made explicit there should not be a need to expose it again. Interviewees tend to resort to unnecessary repetition in an attempt to emphasise the specific messages they intend to deliver. Bull and Mayer (1993) identify repetition of answers to previous questions as one of the ways in which politicians avoid answering questions. The term **soundbite** has recently been associated with this behaviour by politicians during interviews or public speeches. The example in Figure 3.3 shows a case in which essentially the same answer is given by the interviewee to the successive questions by the interviewer[15].

For questions, verbatim repetition often indicates that a satisfactory answer has not been received. Example 1.1 in Chapter 1 illustrates rather well how repetition of a question can be used – a dozen times – by the interviewer to reject an invalid reply. This is made clear by the explicit rejection in turn (11) as shown in Figure 3.4.

[15] This fragment is part of Interview 2 from the corpus study described in Chapter 4. A complete transcript can be found at http://mcs.open.ac.uk/nlg/non-cooperation/.

Turn	Spkr.	Speech
(1)	MILIBAND	These strikes are wrong at a time when negotiations are still going on. But parents and the public have been let down by both sides because the government has acted in a reckless and provocative manner. After today's disruption, I urge both sides to put aside the rhetoric, get round the negotiating table and stop it happening again.
(2)	GREEN	I listened to your speech in Wrexham where you talked about the Labour Party being a movement. A lot of people in that that movement are the people who are on strike today and they'll be looking at you and thinking 'Well, You're describing these strikes as wrong. Why aren't you giving us more leadership as the leader of the Labour movement?'
(3)	MILIBAND	At a time when negotiations are still going on I do belie ve these strikes are wrong. And that's why I say both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and sort the problem out. Because the public and parents have been let down by both sides. The government has acted in a reckless and provocative manner.
(4)	GREEN	I spoke to Francis Maude before I came here and the tone he was striking was a very conciliatory one. Do you think there's a difference between the words they are saying in public and the attitudes they strike in private behind the negotiations? Are the negotiations in good faith would you say?
(5)	MILIBAND	What I say is that the strikes are wrong at a time when negotiations are still going on. But the government has acted in a reckless and provocative manner in the way it has gone about these issues. After today's disruption, I urge both sides to get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.

Figure 3.3: Repetition used by an interviewee for producing soundbites

Complete and Incomplete Replies. As stated above, whether a response to a question constitutes a valid reply depends on the extent to which the information requested is provided. Analysing equivocation in political interviews Bull and Mayer (1993) and Bull (2003) argue that incomplete replies are somewhere between valid and invalid replies, as the information conveyed is relevant but insufficient. In the case of valid multi-barrelled questions for instance, or when the interviewer asks several questions in the same turn, the interviewee can reply to only part of these. In the latter case, individual questions are labelled separately and reply labels would indicate the question they address in their reference. This allows tracking what

Turn	Spkr.	Speech
(1)	Paxman	(interrupting) Did you threaten to overrule him?
(2)	Howard	I, I, was not entitled to instruct Derek Lewis, and I did not instruct him.
(3)	Paxman	Did you threaten to overrule him?
(4)	Howard	The truth of the matter is that Mr. Marriott was not suspended. I-
(5)	Paxman	(overlapping) Did you threaten to overrule him?
(6)	Howard	-did not overrule Derek Lewis.
(7)	Paxman	Did you threaten to overrule him?
(8)	Howard	I have accounted for my decision to dismiss Derek Lewis in great detail, before the House of Commons.
(9)	Paxman	It's a straight yes-or-no question and a straight yes-or-no answer: did you threaten to overrule him?
(10)	Howard	I discussed the matter with Derek Lewis. I gave him the benefit of my opinion. I gave him the benefit of my opinion in strong language, but I did not instruct him because I was not entitled to instruct him. I was entitled to express my opinion and that is what I did.
(11)	Paxman	With respect, that is not answering the question of whether you threatened to overrule him.

Figure 3.4: Repetition used by an interviewer for rejecting replies

questions have been answered and which ones have not. In multi-barrelled questions, on the other hand, several pieces of information are requested by the same action and only one label is used. Providing all necessary information could then take several reply actions, none of which on its own would constitute a complete reply. The completeness flag allows for dealing with this situation at the level of a turn. New, accurate and relevant yet incomplete individual replies are labelled as valid replies with an incomplete flag, until the interviewee has provided for all the information requested, at which point the action is labelled as a valid reply with the complete binary flag. The fragment in Figure 3.5 illustrates this[16]. Incomplete replies are not considered non-cooperative features, provided that at least one reply action in the turn is complete. If none of them are, then one non-cooperative feature is considered at the end of the turn as the interviewee has failed to meet the obligation of providing all the information requested in a valid

[16]This fragment is part of Interview 1 from the corpus study described in Chapter 4. A complete transcript can be found in Appendix A.

question.

Turn	Spkr.	Speech	Action Label
13	BRODIE	This is echoing what George Bush said isn't it about how we will go not just for the perpetrators but for those who harbour him,	**(16)** : `valid-statement`
		and you are talking about the Taliban?	**(17)** : `valid-question` **N**
14	BLAIR	Well, for all those people who have been in a position where they have been helping or harbouring terrorism, the way that it operates, camps that are dedicated to training people in it. These are people trained in these camps who go out and basically wreak havoc wherever they can, killing many, many innocent people. And although what happened last week is obviously an atrocity almost beyond our imagination, it is not an isolated incident, in that sense, there has been a history going back over several years.	**(18)** : `valid-reply@`(17) **I**
		Now you mention the Taliban, the Taliban have a very clear choice, the Taliban either cease to help or harbour those that are fermenting terrorism or they will be treated as part of the terrorist apparatus themselves. Now they have that choice and they should consider very, very carefully the consequences that they face at this moment of choice.	**(19)** : `valid-reply@`(17) **C**

Figure 3.5: Labelling of complete and incomplete valid replies

3.5.2 An Example Interview Annotated with Action Labels

Figure 3.6 shows a (hand-crafted) political interview. The same fragment is annotated with the labels above in Figure 3.7(a). A semi-automatic method for arriving at such annotation is described and evaluated in Chapter 4. Once labelled, a dialogue is represented as a sequence of turns, each one with a speaker and a list of action labels. In the example, the dialogue is represented as shown in Figure 3.7(b) which, formalised as explained below is the input to the method for measuring non-cooperation described in the next chapter.

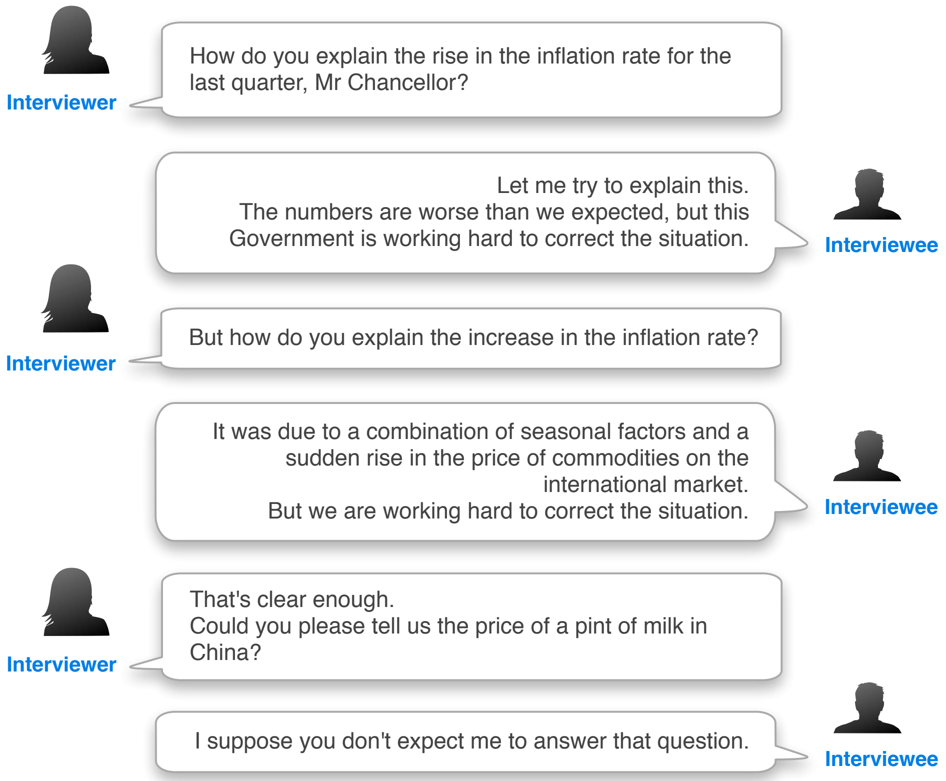


Figure 3.6: A hand-crafted political interview transcript

Formalisation

The action label taxonomy for political interviews is formalised as: $T_{PI} = \{$`valid-statement`, `invalid-statement`, `valid-question`, `invalid-question`, `valid-reply`, `invalid-reply`, `acceptance`, `rejection`$\}$. In a political interview dialogue D_{PI} each turn t_i is a pair (s_i, L_i), where $s_i \in \{\texttt{ir}, \texttt{ie}\}$ is the speaker and L_i is a sequence of instances of action labels from the taxonomy with adequate parameter values (identifier, reference and binary flags as applicable). Two further restrictions apply to correctly labelled dialogues:

- $s_i = \texttt{ir} \Rightarrow \texttt{valid-reply} \notin L_i$
- $s_i = \texttt{ie} \Rightarrow \texttt{valid-question} \notin L_i$

Turn	Spkr.	Speech	Action Label	
1	IR	How do you explain the rise in the inflation rate for the last quarter, Mr Chancellor?	(1) : `valid-question`	**N**
2	IE	Let me explain this.	(2) : `acceptance`@(1)	
		The numbers are worse than we expected, but this Government is working hard to correct the situation.	(3) : `invalid-reply`@(1)	**I**
3	IR	But how do you explain the increase in the inflation rate?	(3) : `valid-question`	**R**
4	IE	It was due to a combination of seasonal factors and a sudden rise in the price of commodities on the international market.	(5) : `valid-reply`@(1)	**C**
		But we are working hard to correct the situation.	(6) : `invalid-reply`@(1)	**C**
5	IR	That's clear enough.	(7) : `acceptance`@(5)	
		Could you please tell us the price of a pint of milk in China?	(8) : `invalid-question`	**N**
6	IE	I suppose you don't expect me to answer that question.	(9) : `rejection`@(8)	

IR: **Interviewer**, IE: **Interviewee**; **N**: New, **R**: Repeated, **C**: Complete, **I**: Incomplete.

(a) Political interview transcript annotated with action labels

Turn	Spkr.	Action Label	
1	IR	(1) : `valid-question`	**N**
2	IE	(2) : `acceptance`@(1)	
		(3) : `invalid-reply`@(1)	**I**
3	IR	(4) : `valid-question`	**R**
4	IE	(5) : `valid-reply`@(1)	**C**
		(6) : `invalid-reply`@(1)	**C**
5	IR	(7) : `acceptance`@(5)	
		(8) : `invalid-question`	**N**
6	IE	(9) : `rejection`@(8)	

IR: **Interviewer**, IE: **Interviewee**;
N: New, **R**: Repeated, **C**: Complete, **I**: Incomplete.

(b) Political interview as a list of turns with action labels

$$\begin{aligned}
D_1 = \langle & (\mathtt{ir}, \langle (1) : \mathtt{valid\text{-}question}\ \mathbf{N}\rangle);\\
& (\mathtt{ie}, \langle (2) : \mathtt{acceptance@(1)}; (3) : \mathtt{invalid\text{-}reply@(1)}\ \mathbf{I}\rangle);\\
& (\mathtt{ir}, \langle (4) : \mathtt{valid\text{-}question}\ \mathbf{R}\rangle);\\
& (\mathtt{ie}, \langle (5) : \mathtt{valid\text{-}reply@(1)}\ \mathbf{C}; (6) : \mathtt{invalid\text{-}reply@(1)}\ \mathbf{C}\rangle);\\
& (\mathtt{ir}, \langle (7) : \mathtt{acceptance@(5)}; (8) : \mathtt{invalid\text{-}question}\ \mathbf{N}\rangle);\\
& (\mathtt{ie}, \langle (9) : \mathtt{rejection@(8)}\rangle)\rangle
\end{aligned}$$

(c) Formalised political interview

Figure 3.7: Action labels in a (hand-crafted) political interview

meaning that there can be no valid replies performed by interviewers, nor valid questions put forward by the interviewee[17]. An example of the formalisation is shown in Figure 3.7(c).

3.5.3 The Dialogue Game

Based on the simplifying assumptions above and on the discussion on political interviews in Chapter 2, **political interviews** can be characterised as follows:

> A political interview is a dialogue with two participants, the interviewer and the interviewee, who make their contributions in turns, with complete utterances and only when they hold the floor – that is, there are no interruptions and no overlaps. The interviewer's role is to elicit information from the interviewee by asking valid questions that can be optionally accompanied by one or more statements setting up the context. The interviewee's role is to provide the information requested by the interviewer.
>
> After a valid question is posed by the interviewer, the interviewee is expected to accept it, either explicitly, or by providing a valid reply. By accepting a question, either valid or invalid, the interviewee is obliged to provide a valid reply. Once a question has been answered, the interviewer is expected to accept it, either explicitly or by moving on to the next question until the interview ends.
>
> In cases when a response is not a valid answer to the question, the interviewer is expected to reject it, either explicitly by indicating how the response fails to provide an answer or by repeating

[17] We are using the mathematical notation loosely here as the sequences of labels do not contain unparameterised labels but instances of the form described in Section 3.5.1.

the question. Conversely, if the interviewer asks an invalid question, the interviewee must reject it, explicitly indicating why the question is not valid or why an answer cannot be provided.

As can be seen from the description above, some of the obligations are inherent to the role of each participant, while others depend on the dynamics of the dialogue and on the actions that participants take. The rules of a **dialogue game for political interviews** are informally expressed as shown in Figure 3.8.

Following the observations by Heritage (1985) and Greatbatch (1988) for news interviews, except for the repetition of questions, we treat the omission of explicit acknowledgements as implicit acceptances. Heritage observed that receipts and assessments are systematically avoided by participants and, for instance, interviewers "display their alignment to prior talk largely by designing next questions so as to tacitly presuppose the truth and adequacy of prior reports or to undermine them." (Heritage, 1985, p. 99). This approach also conforms with the treatment of implicit acceptances as proposed by Kreutel and Matheson (2003a).

Interviews often start with a question from the interviewer, so we assume that at the outset of the dialogue there is an obligation on the interviewer to ask a valid question. In the exceptional cases in which the exchange starts with a statement by the interviewee, there is an obligation for the interviewee to make a valid statement.

The rules of the dialogue game for political interviews are formally represented as $G_{PI} = (Allow_{PI}, Introduce_{PI}, DischargePI)$, defined as shown in Figure 3.9[18].

[18]We use "$*$" as a wildcard character for when the validity of the action is not relevant to the application of the rule.

Dialogue Game for Political Interviews (Informal)

1. The interviewer limits himself or herself to:
   - making valid statements,
   - asking valid questions,
   - accepting a contribution from the interviewee, or
   - rejecting a contribution the interviewee.
2. The interviewee limits himself or herself to:
   - making valid statements,
   - producing valid replies to questions,
   - accepting a statement or question from the interviewer, or
   - rejecting a contribution from the interviewer.

3. If the interviewer makes a valid statement, the interviewee must accept it.
4. If the interviewer asks a valid question, the interviewee must accept it.
5. By accepting a question, the interviewee commits him/herself to providing a valid reply to that question.
6. If the interviewee makes a valid statement, the interviewer must accept it.
7. If the interviewee provides a valid reply to a question, the interviewer is obliged to accept it.
8. By accepting a reply, the interviewer commits him/herself to asking a new valid question.
9. If the interviewer makes an invalid statement, the interviewee must reject it.
10. If the interviewer asks an invalid question, the interviewee must reject it.
11. If the interviewer provides an invalid reply, the interviewee must reject it.
12. If the interviewee makes an invalid statement, the interviewer must reject it.
13. If the interviewee asks an invalid question, the interviewer must reject it.
14. If the interviewee provides an invalid reply, the interviewer must reject it.

15. Repeated (valid and invalid) questions are implicit rejections.
16. Statements (valid and invalid) are implicit acceptances of the contributions in the last turn of the other party.
17. New (valid and invalid) questions are implicit acceptances of the contributions in the last turn of the other party.
18. Replies (valid and invalid) are implicit acceptances of the contributions in the last turn of the other party.

Figure 3.8: An informal dialogue game for political interviews

Dialogue Game for Political Interviews (Formal)

$$G_{PI} = (Allow_{PI}, Introduce_{PI}, Discharge_{PI})$$

where

$$\begin{aligned}
Allow_{PI} = \{ & [\texttt{ir} : \{\texttt{valid-statement}, \texttt{valid-question}, \texttt{acceptance}, \texttt{rejection}\}], && (1)\\
& [\texttt{ie} : \{\texttt{valid-statement}, \texttt{valid-reply}, \texttt{acceptance}, \texttt{rejection}\}]\} && (2)\\
Introduce_{PI} = \{ & [(\texttt{ir}, (\texttt{s}) : \texttt{valid-statement}) \rightsquigarrow (\texttt{ie}, \texttt{acceptance@(s)})], && (3)\\
& [(\texttt{ir}, (\texttt{q}) : \texttt{valid-question}\ \mathbf{N}) \rightsquigarrow (\texttt{ie}, \texttt{acceptance@(q)})], && (4)\\
& [(\texttt{ie}, \texttt{acceptance@(q)}) \rightsquigarrow (\texttt{ie}, \texttt{valid-reply@(q)}\ \mathbf{C})], && (5)\\
& [(\texttt{ie}, (\texttt{s}) : \texttt{valid-statement}) \rightsquigarrow (\texttt{ir}, \texttt{acceptance@(s)})], && (6)\\
& [(\texttt{ie}, (\texttt{r}) : \texttt{valid-reply@(q)}) \rightsquigarrow (\texttt{ir}, \texttt{acceptance@(r)})], && (7)\\
& [(\texttt{ir}, \texttt{acceptance}) \rightsquigarrow (\texttt{ir}, \texttt{valid-question}\ \mathbf{N})], && (8)\\
& [(\texttt{ir}, (\texttt{s}) : \texttt{invalid-statement}) \rightsquigarrow (\texttt{ie}, \texttt{rejection@(s)})], && (9)\\
& [(\texttt{ir}, (\texttt{q}) : \texttt{invalid-question}) \rightsquigarrow (\texttt{ie}, \texttt{rejection@(q)})], && (10)\\
& [(\texttt{ir}, (\texttt{r}) : \texttt{invalid-reply}) \rightsquigarrow (\texttt{ie}, \texttt{rejection@(r)})], && (11)\\
& [(\texttt{ie}, (\texttt{s}) : \texttt{invalid-statement}) \rightsquigarrow (\texttt{ir}, \texttt{rejection@(s)})], && (12)\\
& [(\texttt{ie}, (\texttt{q}) : \texttt{invalid-question}) \rightsquigarrow (\texttt{ir}, \texttt{rejection@(q)})], && (13)\\
& [(\texttt{ie}, (\texttt{r}) : \texttt{invalid-reply}) \rightsquigarrow (\texttt{ir}, \texttt{rejection@(r)})]\} && (14)\\
Discharge_{PI} = \{ & [\texttt{*-question}\ \mathbf{R} \succ \texttt{rejection}], && (15)\\
& [\texttt{*-statement} \succ \texttt{acceptance}], && (16)\\
& [\texttt{*-question}\ \mathbf{N} \succ \texttt{acceptance}], && (17)\\
& [\texttt{*-reply} \succ \texttt{acceptance}]\} && (18)
\end{aligned}$$

Figure 3.9: Formal representation of the dialogue game for political interviews

3.6 Summary

This chapter presented the conceptual framework for the study of non-cooperation in dialogue.

We started by making clear the distinction between cooperation in dialogue at the level of the underlying activity (e.g. working together towards exposing the truth on a certain matter) and cooperation at the conversational level (e.g. asking sensible questions and responding to them adequately). Focusing on the latter, called linguistic cooperation, we gave a generic definition of non-cooperation of a speaker with respect to a conversational setting. By illustrating with a definition of non-cooperation specific to political interviews we identified a number of shortcomings. This led to simplifying assumptions that allowed us to revisit the definition in greater detail by focusing on issues pertaining to core speech acts.

Discourse obligations resulting from the rules specifying global dialogue games were introduced as a conceptual means for tracking the behaviours of dialogue participants and detecting instances of non-cooperation. This led to a revisited, more detailed definition of non-cooperation in dialogue. Under this definition, non-cooperation is the extent to which each dialogue participant has failed to meet their obligations. Failures are aggregated into a single, normalised figure for each participant, which allows comparisons between the participants in the same conversation and also across dialogues and any of their fragments.

The concepts were formalised and illustrated with examples from the political interview setting. The chapter closed with a detailed dialogue game for political interviews which will be used as needed in the examples of the rest of the thesis and in the evaluation of the approach.

Having laid the grounds for the systematic study of non-cooperative behaviour, in the next chapter we present, discuss and evaluate a semi-automatic method for assessing non-cooperation in dialogue. In Chapter 5, the concepts introduced here are incorporated in a model of conversational agents that can exhibit and deal with non-cooperative behaviour.

Chapter 4

Measuring Non-Cooperation in Dialogue

The previous chapter introduced the conceptual framework for studying linguistic non-cooperation in naturally-occurring dialogue and, more specifically, in political interviews. This chapter presents and discusses a semi-automatic empirical method for measuring non-cooperation within this framework. Section 4.1 describes the overall approach and Sections 4.2 and 4.4 explain the method in detail in the context of political interviews. Two corpus studies are described in Sections 4.3 and 4.5, reporting on results for reliability and validity of the measure. Section 4.6 discusses the outcomes in connection with the remainder of the thesis.

4.1 Aims and Overview of the Method

The first aim of the method described here is to obtain, semi-automatically, the degree to which each participant in a dialogue behaved according to the role he or she had in the conversation. This calls for a mechanism to assess a participant's contributions in terms of the expectations for their role in the conversational context specified by the dialogue game introduced in Chapter 3. Ideally, this mechanism should be reliable, i.e. it should be possible for anyone to apply the method to dialogues beyond those analysed in this research. It should also be valid, i.e it should correlate to intuitive notions of cooperation in dialogue in the sense that we give it in the thesis.

The second goal is to inform the study of non-cooperation by means of simulation presented in Chapter 5. This requires that the method be suitable for its incorporation, at least in part, into a model of conversational agents that can exhibit and cope with non-cooperative linguistic behaviour. We will come back to this aspect in Section 4.6.

In short, starting with a transcript of a dialogue and a short description of its context, we arrive at a numeric value for each participant indicating to what extent he or she adhered to the rules of the dialogue game. We call this value the **degree of cooperation** of each dialogue participant with respect to a dialogue game.

The measure is obtained in two steps. In the first step, the dialogue transcript is **manually annotated**, identifying a set of features in the participants' contributions. In the second step, these features are **automatically assessed for cooperation** with respect to the dialogue game. Figure 4.1 shows an overview of this process. The icons on the upper-left corner of the boxes for each step indicate whether they are automatic or involve human intervention. The chess board icon inside the box Assessment of Cooperation shows that the dialogue game plays a role in this step. We will use this convention when describing the method in further detail.

Section 4.2 focuses on the annotation step applied to political interviews. Annotations are done in two stages:

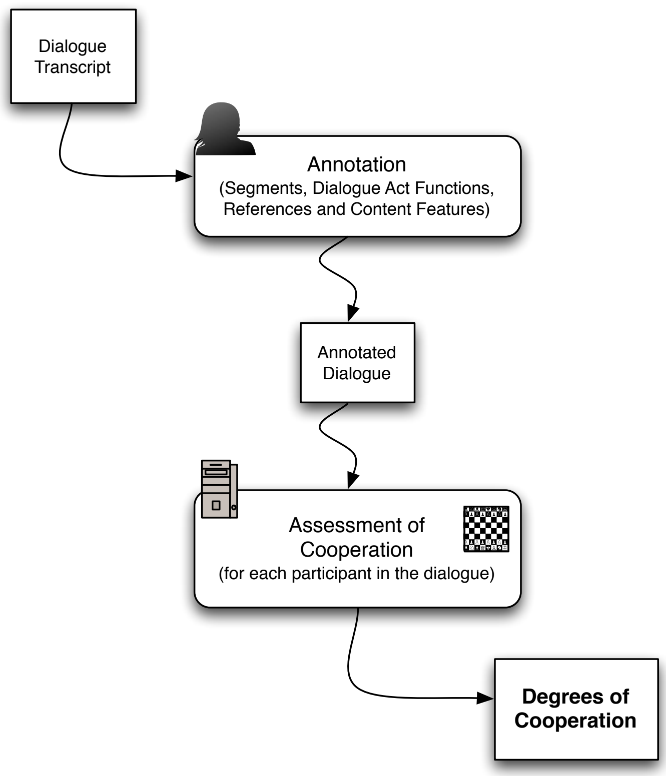


Figure 4.1: Two-step semi-automatic measure of cooperation (overview)

First Stage: each turn in the transcript is divided into **segments** and each segment is assigned a **dialogue act function** and, when applicable, the **segment it responds to**.

Second Stage: each segment identified in the first stage is annotated with **content features**, i.e. qualitative judgements on its content with respect to the context of to the dialogue and to the dialogue history.

The concepts above will be defined precisely in Section 4.2. The outcome of the annotation step is a dialogue in which the contributions of the participants are identified and categorised according to their function and to the qualitative properties of their content. These annotations are then used in the second step to determine the degree of cooperation of each dialogue participant.

Section 4.3 presents a corpus study on the **reliability** of the method. The study involved applying the annotation scheme to six political interviews. Inter-annotator agreement was measured for each annotation stage, including segmentation agreement using Krippendorff's α_U coefficient (Krippendorff, 1995), which was adapted for segmentation of transcribed dialogue (see Section 4.3.1 for details). Agreement was also measured on the assignment of dialogue act functions, referent segments (first annotation stage) and content features (second annotation stage) using multi-rater versions of Cohen's κ (Cohen, 1960; Davies and Fleiss, 1982), Scott's π (Scott, 1955; Fleiss, 1971) and Krippendorff's α (Krippendorff, 2004). After the first stage, the data from individual annotators were aggregated automatically to produce a single annotation for each dialogue. These were used as input for the second annotation stage.

Section 4.4 describes the automatic assessment step applied to political interviews. This involves analysing the annotated dialogues in order to identify departures from the linguistic behaviour that is expected of each participant according to the dialogue game. The assessment process starts by mapping the annotations to **action labels** and computing the **discourse obligations** introduced by these actions, as described in Chapter 3. Next, these action labels and discourse obligations are used to detect **cooperative** and **non-cooperative features**. These are instances in which an action respectively met or failed to meet a discourse obligation. Finally, the degree of cooperation is calculated for each speaker as the ratio of cooperative features over the total number of features (cooperative and non-cooperative). This gives a numerical value between 0 and 1, that is, between **total non-cooperation** where none of the speaker's actions met the obligations prescribed by the dialogue game and **total cooperation** where all of the speaker's actions met their obligations. The concepts above are defined and formalised in Section 4.4, where we also present and discuss the algorithms and data structures involved in each of the computations. Figure 4.2 shows a more detailed schematic representation of the method with the annotation stages and processes just introduced.

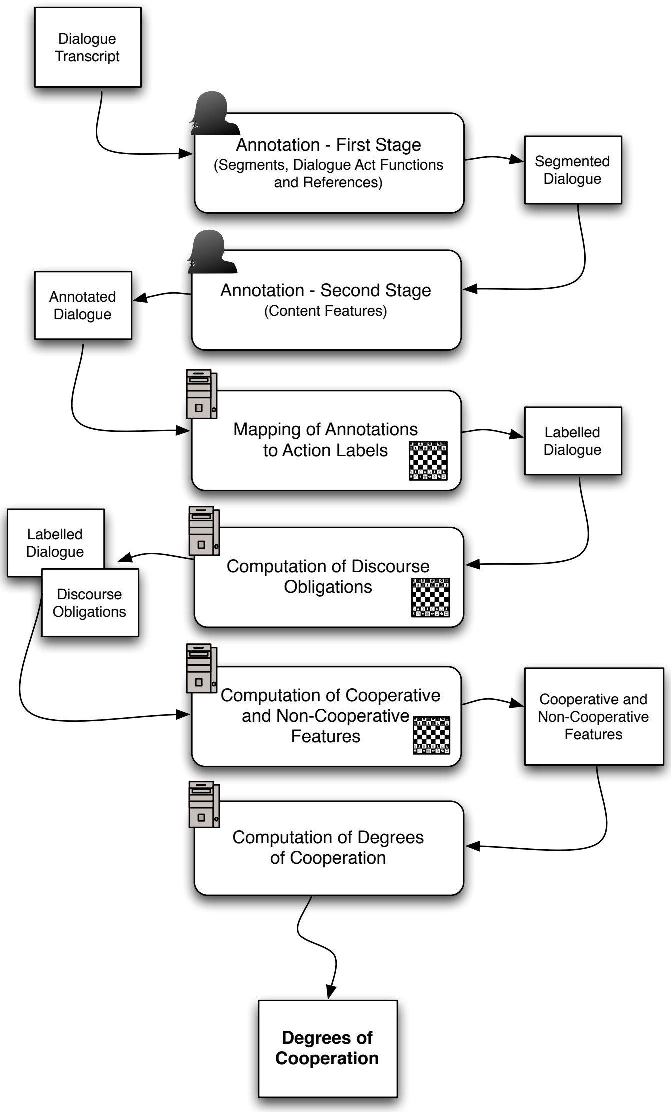


Figure 4.2: Stages of the semi-automatic measure of cooperation

Section 4.5 presents a study on the **validity** of the method. The study involved eliciting human judgement on the behaviour of dialogue participants via an on-line survey. This was performed independently from the previous study and on the same set of dialogue transcripts. Validity of the measure was assessed by testing the correlation between these judgements and the degree of cooperation obtained when applying the method to the aggregated annotations from the second stage. Figure 4.3 shows an overview of the evaluation approach.

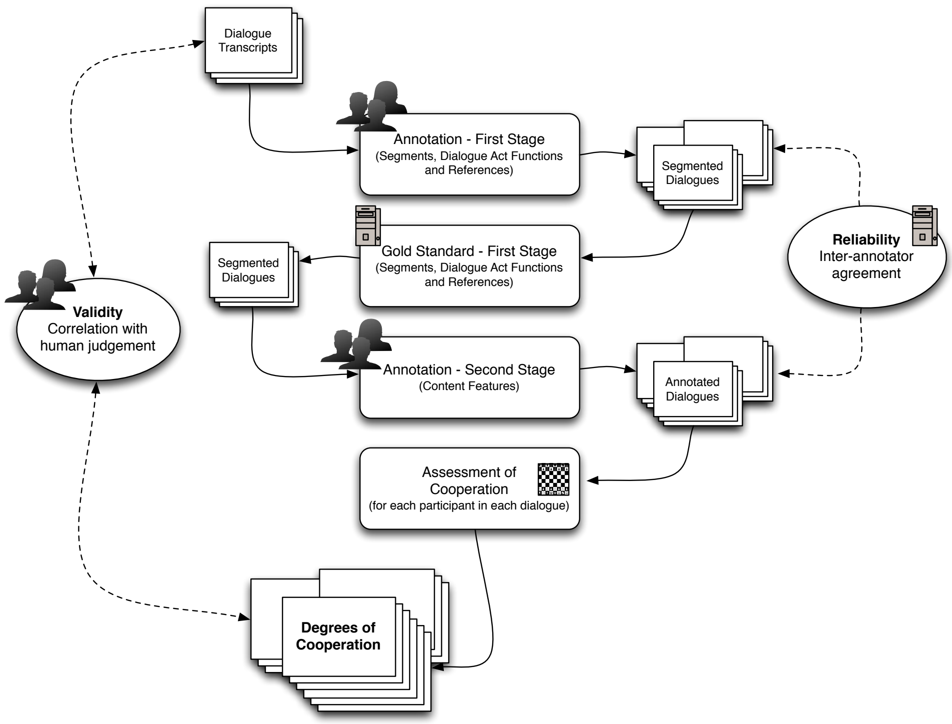


Figure 4.3: Overview of the evaluation approach

4.2 Annotating Political Interviews for the Assessment of Cooperation

This section describes an annotation scheme for naturally-occurring political interviews towards the automatic assessment of conversational cooperation.

The assessment requires that participants' contributions be classified according to their effects and to the quality of their content. Typical examples include distinguishing questions from responses and, further, questions that request for information relevant to the topical context of the interview from those that do not, or responses that provide a meaningful answer to a question from those that fail to do so.

The distinctions pertaining to the content of what participants contribute to the conversation pose a challenge with respect to previous research in dialogue annotation. Classical annotation schemes for dialogue moves (Carletta et al., 1997; Allen and Core, 1997; Poesio et al., 1999) and also those specifically concerned with non-cooperation (Davies, 1997; Cavicchio, 2010), were developed in the context of tightly constrained task-oriented domains. Given the nature of such dialogues, either in human–human interaction analysis or human–computer dialogue modelling, these proposals assume dialogue participants work together towards the resolution of the task at hand. As discussed in Chapter 3, political interviews share some characteristic with task-oriented dialogues (e.g. the roles of the participants are clearly defined). However, it is often the case that politicians and interviewers have contradicting goals, which results in contributions not always meeting the expectations for the conversational context or for the role of the participant in the dialogue.

More recently, annotation schemes have been proposed for other types of dialogue, including fictive expository dialogues (Stoyanchev and Piwek, 2010a), argumentation and persuasion (Georgila et al., 2011), and open-context, multi-modal, multi-party interactions (Bunt, 2009; Bunt, 2011; Bunt et al., 2012). Some of these schemes consider interactions in which participants' goals do not point in the same direction (e.g. the annotation scheme for argumentation and persuasion proposed by Georgila et al.), but to our knowledge this is the first time a coding scheme is proposed for naturally-occurring open-domain political interviews aimed at investigating aspects of linguistic non-cooperation. Political interviews are in a way expository dialogues: the ultimate goal of an interview is to make information publicly available. On the other hand, they differ with respect to the specific roles, both conversational and public, of the participants. There is not, for instance, a clear divide between an expert and a naive participant with respect to the information being elicited, and therefore an answer from the politician can be challenged by the interviewee based on facts. Political interviews are also instances of naturally-occurring dialogue, what differentiates them further form fictive dialogue of the kind studied, for instance, by Stoyanchev and Piwek (2010a).

We define our annotation scheme taking elements from the ISO standard proposed by Bunt et al. (2009; 2010; 2012) and incorporating insights from the analysis of equivocation in political communication proposed by Bull and Mayer (1993) and developed further by Bull (1994; 2003).

4.2.1 A Discussion on Related Work

Previous research on dialogue annotation for non-cooperation is scarce. The only instances of complete research we know of are those of Davies (1997; 2006) and Cavicchio (2010), both in the context of task-oriented dialogues, and more specifically in the HCRC Map Task domain (Anderson et al., 1991; Carletta et al., 1997)[1]. Davies (1994; 1997; 2006) proposes a direct approach

[1]At the moment of writing this thesis, two short papers were published by Asher et al. (2012) and Afantenos et al. (2012) from the STAC project which focuses on analysis and modelling strategic conversation with a strong emphasis on non-cooperation. The papers report on ongoing data collection and preliminary annotation of negotiation dialogues surrounding a board game, but mention is made of future work exploring data from

to annotating cooperation for analysing its relation with effort and task success. Her notion of cooperation is a re-interpretation of Grice's (1975) Cooperative Principle in terms of the risk/effort trade-off of a speaker's dialogue acts and their alignment with the hearer's goals. Her annotation approach shares some characteristics with ours, but cooperation is judged directly by the annotators, as "positive codings (i.e. finding an instance of the behaviour in an utterance), and negative codings (i.e. finding an instance where we believe a particular behaviour should have been used)" (Davies, 2006, p. 43). In addition, codings are weighed towards the rating of cooperation according the their level of effort (low-effort behaviours have lower positive weightings but higher negative weightings, while the converse holds for high-effort behaviours). In her doctoral thesis, Cavicchio (2010) applies Davies' coding scheme to a multi-modal corpus of the Map Task domain and studies the relation between (non-)cooperation and emotions. Her focus is not however on how to assess cooperation in dialogue, but on to what extent psychophysiological indicators of emotion (e.g. heartrate and facial expressions) correlate with cooperative behaviour.

We considered a series of flat annotation schemes[2] similar to Davies' in previous investigations (Plüss, 2010), but abandoned them in favour of the multi-dimensional one presented here for several reasons. First, the open-domain of naturally-occurring political interviews makes the annotation task inherently more complicated and cognitively demanding than in

political debating dialogues. Their modelling approach seems to share the concept of dialogue game with the one presented here, while their approach to dialogue management uses a statistical planning method (reinforcement learning) for acquiring optimal dialogue policies (Rieser and Lemon, 2011), diverging from what we present in Chapter 5.

[2]By flat annotation scheme we mean that one label is chosen for each annotation unit, as shown in Figure 4.4 (Neutral Question, Loaded Question, Reply, Unsolicited Comment, etc.), as opposed to a multi-dimensional coding scheme in which each unit receives several content feature labels that depend hierarchically on the dialogue act annotations of the first stage.

controlled domains such as the Map Task. This difficulty was worsened by a flat taxonomy which made annotating each contribution considerably harder[3]. Although preliminary annotation schemes are not discussed here in further detail, for illustration we point the reader to Figure 4.4 which shows a decision tree intended at guiding the annotation using one of such schemes.

The second reason is that schemes like Davies' already include the normative notion of dialogue game we use later in the assessment of cooperation. This reduces the flexibility of the coding scheme, as the assessment of cooperation is part of the annotation process. By detaching these steps, the method proposed here allows for assessment of cooperation of the same annotated data using different dialogue games, e.g. to explore how the same behaviour would be perceived by audiences with different cultural backgrounds.

The third reason for favouring the coding approach described below is that it results in a classification of contributions that is closer to the structure of the dialogue game presented in Chapter 3 and also to the model of conversational agents discussed in Chapter 5. This allows for the insights obtained by empirical analysis to be easily incorporated in the model of conversational agents.

Dividing the annotation in two stages and dealing with segmentation and dialogue act function annotation separately from the qualitative analysis of the content had several advantages. First, it made individual annotations considerably easier and less time consuming by reducing the cognitive load on the annotators. Second, it draws us closer to previous work on segmentation and dialogue act coding. By building on previous research for the first

[3] A flat annotation scheme was a first natural choice as these generally score better in inter-annotator agreement than multi-dimentional ones, as pointed out by Artstein and Poesio (2008, p. 33).

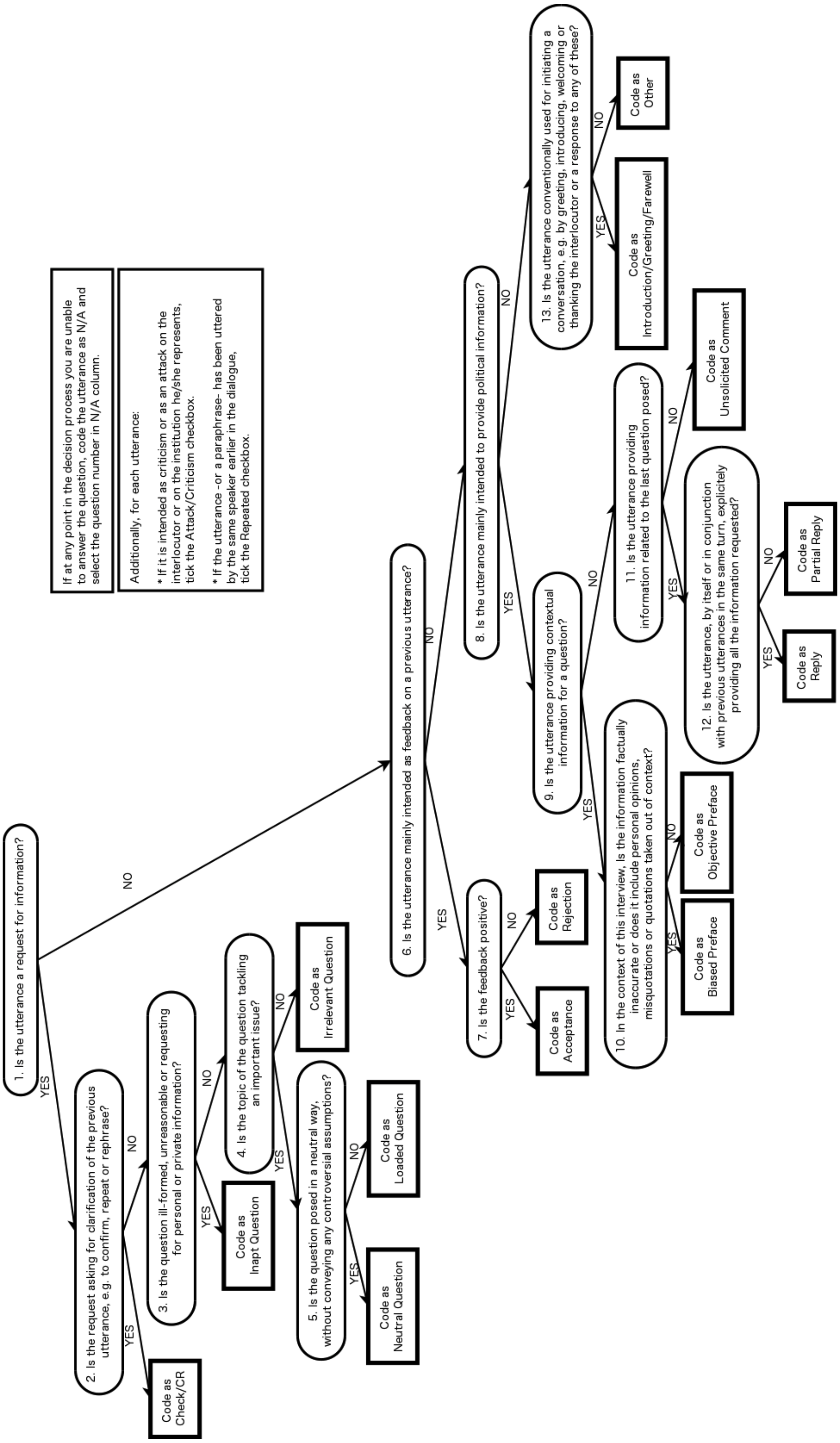


Figure 4.4: Flat coding scheme from preliminary investigations

stage, we could focus on the more specific and novel qualitative judgements on the content of the contributions in the second stage.

For the first stage, we followed the recommendations put forward in the ISO standard proposal by Bunt et al. (2009; 2010; 2012), simplifying the terminology and some aspects of the scheme when needed, drawing from work by Carletta et al. (1997), Allen and Core's (1997) DAMSL, Traum and Hinkelman's (1992) Conversation Acts theory – following Poesio and Traum (1997; 1998) and proposed as a standard by the Discourse Resource Initiative (Initiative, 1997) [4] – and Stoyanchev and Piwek (2010a; 2010b).

For the second stage, we identified a set of dimensions on which the content of a contribution are judged. These are based, in part, on Bull and Mayer's (1993) and Bull's (1994; 2003) extensive work on the micro-analysis of equivocation in political discourse.

4.2.2 First Annotation Stage: Segmenting Turns and Annotating Dialogue Acts and Referent Segments

As mentioned above, in the first annotation stage the turns in dialogue transcripts are divided into **segments** and each segment is assigned a **dialogue act function** and, when applicable, the **segment it responds to**. Below we define these concepts and the annotation scheme. We also describe the annotation procedure and provide examples[5].

[4] It is worth noting that in DAMSL (Allen and Core, 1997), Conversation Acts theory (Traum and Hinkelman, 1992) and the ISO standard proposed by Bunt et al. (2012) the same utterance can be annotated with several dialogue act functions operating at different levels (e.g. turn-taking, grounding, core speech acts). As discussed in Chapter 3, our focus is on core speech acts, so we will annotate each segment with a single dialogue act function. This however does not make our approach incompatible with the aforementioned as noted in Section 4.6.

[5] Copyright of all interview fragments belong to the respective broadcasting company and are reproduced here for the purpose of examination towards an academic degree.

Definitions

Turn: a speaker's continued contribution before the other dialogue participant takes over. In the transcript, this is the fragment of text next to a speaker label – i.e. `IR` (interviewer) or `IE` (interviewee).

Segment: the longest stretch (i.e. a continuous substring of the transcript) of a turn, on the same topic (see further details regarding topicality below), that can be labelled with a single dialogue act function. Stretches of a turn can belong to only one segment – i.e. segments do not overlap – and some stretches can remain unannotated.

Dialogue Act Function: the conversational action performed by a segment. Dialogue acts functions can be **responsive** or **initiating**[6], depending on whether they initiate an exchange pair or respond to an initiation. Typical examples are questions (initiating) and their replies (responsive).

Referent Segment: a segment in a previous turn of the other speaker to which the current segment responds[7]. By definition, every segment with a responsive dialogue act function must have an associated referent segment. Conversely, segments with an initiating dialogue act function do not have a referent segment.

[6]The distinction between responsive and initiating dialogue act functions is analogous to that between **backward-looking** and **forward-looking** functions in DAMSL (Allen and Core, 1997), or to the distinction between dialogue acts with and without a **functional dependence link** in the ISO standard proposed by Bunt et al. (2012).

[7]Referent segments are called **functional dependence relations** in the ISO standard proposed by Bunt et al. (2012). The name stems from the observation that "the semantic content of these types of dialogue act depend crucially on which previous dialogue act they respond to" (Bunt et al., 2012, p. 432).

Annotation Instructions

The instructions for segmenting and annotating a political interview in the first stage are summarised as follows:

1. For each turn in the dialogue:

   (a) Segment the turn by selecting the stretches of speech that have a clear dialogue act function.

   (b) Assign a dialogue act function to each segment, identifying whether the dialogue act is initiating an exchange (i.e. requesting information, giving information as context for an upcoming question, etc.), or responding to a previous dialogue act (i.e. accepting a question or an answer, answering a question, rejecting a premise, providing additional information, etc.).

   (c) For each responsive segment, select the segment that caused the response.

Dialogue Act Taxonomy

As said, dialogue acts are the actions speakers perform in a conversation. Chapter 3 described political interviews as a subtype of information-seeking dialogues, usually structured as a sequence of question–answer pairs, in which one of the participants asks the questions and the other provides the answers. Questions are sometimes preceded by a few statements setting up the context or with an observation on the previous answer. Similarly, answers can be preceded or replaced by remarks on the previous question.

When identifying these actions, focus should be on the *function* they play in the dialogue, rather than, for instance, on their syntactic form[8]. So,

[8] This is a generalisation of Bull's (1994; 2003) *functional criterion* for identifying questions in political interviews.

for example, a question needs not necessarily be in interrogative form to function as a request for information. Similarly, a rhetorical question can be conveying information rather than asking for a reply.

We consider two main classes of functions for dialogue acts: **initiating** and **responsive**. Initiating dialogue acts are primarily meant to provoke a response by the other speaker – as opposed to being themselves responses to previous dialogue acts. Responsive dialogue acts are mainly reactions of the speaker to a previous (initiating or responsive) action of the other party.

Initiating dialogue acts are further divided into **information giving** and **information requesting** dialogue acts. For the annotation, we refer to these as Init-Inform and Init-InfoReq, respectively:

- **Init-Inform** dialogue acts have as main function to make a piece of information (e.g. a fact, an opinion) available to the hearer.
- **Init-InfoReq** dialogue acts are aimed at requesting a piece of information from the hearer.

Responsive dialogue acts are further divided into **information giving**, **accepting** and **rejecting** dialogue acts. For the annotation, we refer to these as Resp-Inform, Resp-Accept, Resp-Reject, respectively:

- **Resp-Inform** dialogue acts have as main function to make a piece of information (e.g. a fact, an opinion) available to the hearer in response to a previous contribution.
- **Resp-Accept** dialogue acts are mainly aimed at indicating that the speaker is satisfied with a previous contribution of the other party (positive feedback).

– **Resp-Reject** dialogue acts have as principal role indicating that the speaker objects to the contribution of the other party (negative feedback).

Figure 4.5 shows the dialogue act taxonomy.

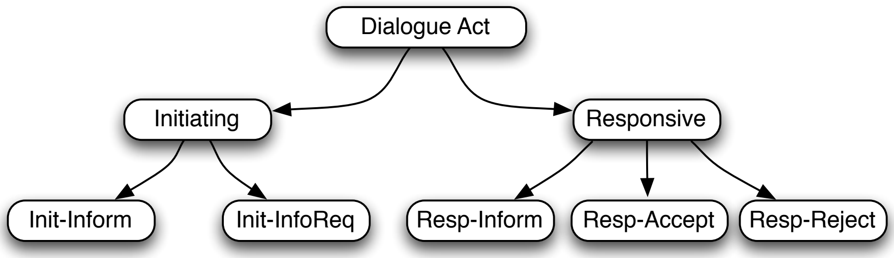


Figure 4.5: Dialogue Act Taxonomy

Deciding What Constitutes a Segment

When choosing the stretches of a turn that constitute separate segments two criteria are followed:

1. The stretch has to be of a length such that it can be assigned one of the available dialogue act functions, and

2. its contents have to request for or convey a clearly identifiable, ideally unique piece of information, or several pieces of information on the same topic.

Example 4.1. Consider the following turn from the interview between BBC presenter Jeremy Paxman and then Home Secretary Michael Howard first introduced in Example 1.1:

INTERVIEWER Right, uh... can you help us with this then? You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it. Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me. Are you saying Mr Lewis is lying?

(*Newsnight*, BBC, 1997)

The turn contains two questions and three quotations from two different sources. The first question is in fact an invitation to comment on an issue, i.e. a politeness formula with a function that does not match any of the options listed above. The quotations are setting up the context for the request for information that comes at the end of the turn. This turn is thus segmented as follows[9]:

INTERVIEWER Right, uh... can you help us with this then? **(1.1)**[You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it.] **(1.2)**[Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me.] **(1.3)**[Are you saying Mr Lewis is lying?]

When segmenting requests for information we distinguish between long single-barrelled questions and multi-barrelled questions. A single-barrelled question asks for one piece of information or for several pieces of the same

[9]Note that the stretch "Right, uh... can you help us with this then?" is not assigned to any segments.

kind of information (e.g. a confirmation, an opinion or view on a certain matter, the name of one or more people, etc.) and should belong in one segment. Multi-barrelled questions, on the other hand, are sequences of separate questions posed together which are assigned to separate segments.

Example 4.2. Consider the following fragment from an interview between CBS Evening News anchor Dan Rather and then American vice-president George H. Bush during the coverage of the 1988 presidential campaign:

INTERVIEWEE	(Interrupting) I wanted those hostages. I wanted Mr Buckley out of there-
INTERVIEWER	(Interrupting) But you made us hypocrites in the face of the world. How could you sign on to such a policy? And the question is what does that tell us about your record?

(*CBS Evening News*, CBS Corporation, 1988)

In the second turn, the interviewer starts with a response to the contribution of the interviewee and continues with a two-barrelled question. Segmentation of this turn is as follows:

INTERVIEWER	(Interrupting) **(2.1)**[But you made us hypocrites in the face of the world.] **(2.2)**[How could you sign on to such a policy?] **(2.3)**[And the question is what does that tell us about your record?]

Similarly, long responses are segmented identifying the stretches of speech that can be assigned a unique dialogue act function. If the function is to provide information, then pieces of information on different topics should belong in separate segments.

Example 4.3. Consider the following turn as a continuation from the previous example:

INTERVIEWEE	The same reason the President signed on to it. When a CIA agent is being tortured to death, maybe you err on the side of a human life. But everybody's admitted mistakes. I've admitted mistakes. And you want to dwell on them, and I want to talk about the values we believe in and experience and the integrity that goes with all of this, and what's I'm going to do about education, and you're, there's nothing new here. I thought this was a news program. What is new?

(*CBS Evening News*, CBS Corporation, 1988)

According to the criterion above, this turn is segmented as:

INTERVIEWEE	**(3.1)**[The same reason the President signed on to it. When a CIA agent is being tortured to death, maybe you err on the side of a human life.] **(3.2)**[But everybody's admitted mistakes. I've admitted mistakes. And you want to dwell on them,] **(3.3)**[and I want to talk about the values we believe in and experience and the integrity that goes with all of this, and what's I'm going to do about education], and you're, **(3.4)**[there's nothing new here. I thought this was a news program. What is new?]

The rest of the section shows how each of the segments in these examples are annotated with dialogue act functions and referent segments.

Selecting a Dialogue Act Function

The first decision when selecting a dialogue act function is whether it is initiating or responsive. Annotators should ask themselves the question:

- *Can I identify a segment to which this one responds?*

If the answer is 'No', then the segment is initiating. Otherwise, it is responsive.

Exceptions. Follow-up questions and clarification requests are exceptions to this rule. Although follow-up questions refer to previous contributions, as they also have an initiating function we favour this aspect and regard them as initiating dialogue acts. Clarification requests are left unannotated.

Selecting an Initiating Dialogue Act Function

Once they have decided that a segment is initiating, annotators should ask themselves the following question:

- *Is the segment only aimed at providing information or is it requesting a contribution from the other party?*

In the first case, the segment is annotated as **Init-Inform**. In the second case, it is annotated as **Init-InfoReq**. Going back to Example 4.1, the segments are annotated as follows:

INTERVIEWER	**(1.1)**	You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it.	**Init-Inform**
	(1.2)	Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me.	**Init-Inform**
	(1.3)	Are you saying Mr Lewis is lying?	**Init-InfoReq**

Example 4.4. As a further example, consider the following two turns from an interview between BBC presenter Caron Keating and UK Prime Minister Margaret Thatcher in relation with Kampuchea (now Cambodia):

INTERVIEWER	Although Pol Pot is actually on the border at the moment, it said only in Thursday's paper that he is actually there.
INTERVIEWEE	Yes, indeed. And, of course,...

(*Blue Peter*, BBC, 1988)

Although the first turn is in the form of a statement, it is inviting a response from the interviewee. In these cases, it is helpful to bear in mind the specific roles of interlocutors in an interview. Noting that this is said by the interviewer is a good indicator that it is primarily about eliciting a response. The annotation is thus as follows:

INTERVIEWER	**(4.1)**	Although Pol Pot is actually on the border at the moment, it said only in Thursday's paper that he is actually there.	**Init-InfoReq**

Selecting a Responsive Dialogue Act Function

Once they have decided that a segment is responsive, annotators should ask themselves the following question:

- *Is the segment meant as providing feedback on or an assessment of a previous contribution or is it aimed at making a new piece of information available to the other party?*

In the first case, the segment is annotated as **Resp-Accept** or **Resp-Reject**, depending on whether the feedback or assessment is positive or negative. In the second case, it is annotated as **Resp-Inform**.

If a segment is an explicit acceptance of the previous contribution it is annotated as **Resp-Accept**. For example, if after a *wh*-question (i.e. what, when, where, which, who, how, etc.) the interviewee starts his response with "*Okay*", this could be considered an acceptance and not, say, a reply to a *yes/no*-question. This, however, would depend on the rest of the response. Other statements like "*That is a very good question*" are also acceptances. After responses, expressions like "*Thanks*" or "*Right*" usually constitute acceptances. Also, more explicit cases like "*Well, that answers my question*".

If a segment is an objection to a previous contribution it is annotated as **Resp-Reject**. For example, if after an alternative or disjunctive question (i.e. those in which two or more alternatives are presented for the hearer to choose from), the interviewee starts his response with "*No*" this is considered a rejection (and not, say, a reply to a *yes/no*-question). Statements like "*I will not answer that question*" are also rejections. Although this depends heavily on the rest of the contribution, after responses, an utterance like "*Excuse me*" might constitute a rejection. Also, more explicit cases like "*You are not answering the question*".

Exceptions. A special case are responses like "*I do not have an answer for that question*" or "*We will only know in due time*". As they express the inability of the speaker to provide an answer, they are considered informative responses, as opposed to rejections, and should be annotated as **Resp-Inform**.

Selecting a Referent Segment

In cases in which the current segment refers to more than one previous segment (e.g. acceptances and rejections of long contributions), annotators should choose the last segment of the set (i.e. the most recent one).

Going back to Examples 4.2 and 4.3, the segments are annotated as follows (we use the notation "**@(<segment-number>)**" to indicate referent segments and shorten the speaker labels as IR and IE for interviewer and interviewee, respectively):

IR	**(2.1)**	But you made us hypocrites in the face of the world.	**Init-Inform**	
	(2.2)	How could you sign on to such a policy?	**Init-InfoReq**	
	(2.3)	what does that tell us about your record?	**Init-InfoReq**	
IE	**(3.1)**	The same reason the President signed on to it. When a CIA agent is being tortured to death, maybe you err on the side of a human life.	**Resp-Inform**	**@(2.2)**
	(3.2)	But everybody's admitted mistakes. I've admitted mistakes. But you want to dwell on them,	**Resp-Inform**	**@(2.3)**
	(3.3)	I want to talk about the values we believe in and experience and the integrity that goes with all of this, and what's I'm going to do about education	**Resp-Inform**	**@(2.3)**
	(3.4)	there's nothing new here. I thought this was a news program. What is new?	**Resp-Reject**	**@(2.3)**

4.2.3 Second Annotation Stage: Selecting Content Features

In the second annotation stage, segmented dialogues are annotated with **content features**. Below we define this concept and give a taxonomy. We also describe the annotation procedure and provide several examples.

Definitions

Content Features: a set of qualitative binary judgements on the content of a segment.

Content Dimensions: the different aspects on which the content of a segment is judged. Dimensions are determined by the dialogue act function assigned to the segment in the previous annotation stage.

Annotation Instructions

In this annotation stage, an annotator receives a dialogue in which the turns have been segmented and annotated with dialogue act functions from the taxonomy described above and, when applicable, with referent segments. The segmented dialogue is accompanied by a brief description of the context in which the interview took place and the main topic(s) discussed.

The instruction for annotating the content features in these dialogues are summarised as follows:

1. Read the context of the interview.

2. For each turn:

    (a) Judge the content of each annotated segment in the dimensions given for the associated dialogue act function following the guidelines. In doing so, identify objective quotations, neutral and relevant questions, complete answers, controversial statements, misquotations, ill-formed or loaded questions, incomplete answers, irrelevant comments, etc.

3. Once the interview is annotated, review each segment and check that the judgement on the content features has not changed while annotating further turns. If it has changed, adjust the values accordingly.

Content Feature Taxonomy

According to the definition above, the content features of a segment are a set of binary qualitative judgments on its content with respect to the context of the interview and to what the participants have contributed so far. The number of judgements corresponds to a set of dimensions (e.g. topicality, relevance, accuracy) associated with each dialogue act function in the taxonomy presented in the previous section.

In the rest of this section, we describe the dimensions for each dialogue act function, except for **Resp-Accept** and **Resp-Reject** that have no associated content features. The content feature taxonomy is shown in Figure 4.6 as an extension of the dialogue act taxonomy presented in Figure 4.5.

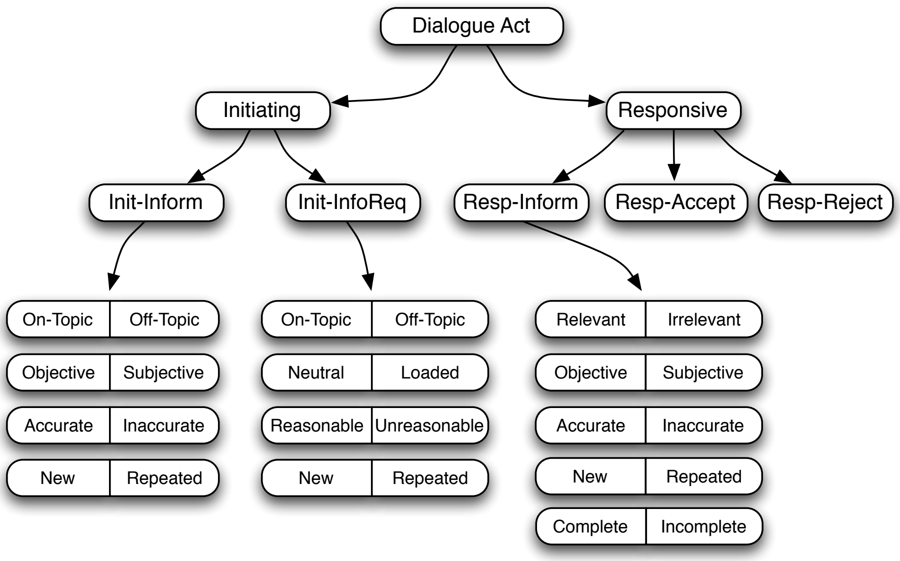


Figure 4.6: Content Feature Taxonomy

Content Features for **Init-Inform** Segments

For segments annotated with an **Init-Inform** dialogue act function we consider the following dimensions:

On-Topic | Off-Topic: whether or not the information provided in the segment is related to the subject matter of the interview.

Objective | Subjective: whether the information provided is unbiased, impartial and evidence-based or conveys the opinion or point of view of the speaker.

Accurate | Inaccurate: whether the information provided is correct or presents imprecisions, errors or false statements.

New | Repeated: whether the information provided is novel or has been provided earlier in the dialogue by the same speaker.

Content Features for Init-InfoReq Segments

For segments annotated with an **Init-InfoReq** dialogue act function we consider the following dimensions:

On-Topic | Off-Topic: whether or not the information requested in the segment is related to the subject matter of the interview.

Neutral | Loaded: whether the request for information is straight-forward and non-leading or contains controversial assumptions, bias, criticisms or accusations.

Reasonable | Unreasonable: whether the information requested is available to the hearer (bearing in mind his public role, common sense, etc.) or it is not expected that he or she would be able to provide it.

New | Repeated: whether the information requested is novel or has been requested earlier in the dialogue by the same speaker.

Content Features for **Resp-Inform** Segments

For segments annotated with an **Resp-Inform** dialogue act function we consider the following dimensions:

Relevant | Irrelevant: whether or not the information provided in the current segment was requested in the segment to which it responds.

Objective | Subjective: whether the information provided is unbiased, impartial and evidence-based or conveys the opinion or point of view of the speaker.

Accurate | Inaccurate: whether the information provided is correct or presents imprecisions, errors or false statements.

New | Repeated: whether or the information provided is novel or has been provided earlier in the dialogue by the same speaker.

Complete | Incomplete: whether the information given in this segment satisfies the information requested in the segment to which it responds or there are still pieces of the information requested that have not been provided.

A Discussion on the Choice of Dimensions

The dimensions above were selected with the aim of obtaining a qualitative profile of each contribution a participant makes in an interview. As far as possible, we wanted these judgements to be independent from the normative constraints of the dialogue game for political interviews presented in Chapter 3 and from the roles of the dialogue participants. This follows a principle of modularity in view of the incorporation of the method in the model of conversational agents discussed in Chapter 5. Also, as mentioned earlier, it allows for data annotated with this scheme to be analysed by considering variations of the dialogue game, for instance, to investigate how cooperation is perceived by in different cultural backgrounds.

In their analysis of equivocation in political interviews, Bull and Mayer (Bull and Mayer, 1993) categorise the ways in which politicians fail to answer questions, proposing a hierarchy of 11 categories and 32 subcategories[10]. Although we were not aiming for such a fine-grained granularity in our study, their insights informed our choice of dimensions, while keeping in sight the requirement for independence of the coding scheme with respect to a particular dialogue game.

Selecting Content Features

When judging the content of a segment annotators must consider, to the best of their knowledge, several elements of the context of the conversation (e.g. topical, political, historical), as well as common sense, world knowledge, etc. They must also take into account previous contributions of both participants, and in some cases contributions made later on in the dialogue. Every time annotators make a judgement, they should ask themselves the following question:

- *Do I have any evidence to make this choice?*

If the answer is 'Yes', then they can go ahead with their choice. Otherwise, they must **be charitable**. This means that, for instance, if is not possible to determine whether the information provided in a segment is accurate or not, the first option is chosen. Similarly, if whether a question is reasonable or not cannot be decided, then it is is considered reasonable.

[10] This work was developed further by Bull (Bull, 1994; Bull, 2003) to include questions, replies and non-replies in political interviews.

The rest of the section, provides a few examples and indicate how their content features are annotated.

Selecting Init-Inform Content Features

Consider the following interview context:

> *"BBC presenter Jeremy Paxman questions former UK Home Secretary Michael Howard with respect to a meeting in 1995 between Howard and the head of the Prison Service, Derek Lewis, about the dismissal of the governor of Parkhurst Prison, John Marriott, due to repeated security failures. The case was given considerable attention in the media, as a result of accusations by Lewis that Howard had instructed him, thus exceeding the powers of his office."*

This is the context of Example 4.1 given in the previous section. The first two segments of the example were annotated as follows:

INTERVIEWER	**(1.1)**	You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, I threatened to instruct him to do it.	**Init-Inform**
	(1.2)	Derek Lewis says Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me.	**Init-Inform**

As we pointed out earlier, the speaker is presenting two literal quotations setting the context for an upcoming question. There is no evidence that the quotations are false or erroneous and they have not been mentioned before (this is the beginning of the interview fragment). Therefore, for both segments the following (underlined) content features are selected:

On-Topic	**Off-Topic**
Objective	**Subjective**
Accurate	**Inaccurate**
New	**Repeated**

If, later on, the interviewee noted, for instance, that the quotations are inaccurate and we have reasons to trust his argument, then the third judgement would be reviewed.

Example 4.5. Now, consider the following segment, a few turns later in the same interview:

INTERVIEWER	**(5.1)**	Mr Lewis says, If I did not change my mind and suspend Marriot he would have to consider overruling me.	**Init-Inform**

This is another quote with essentially the same information conveyed by segment **(1.2)** above. The selection of features in this case is as follows:

On-Topic	**Off-Topic**
Objective	**Subjective**
Accurate	**Inaccurate**
New	**Repeated**

The context for the fragments presented in Examples 4.2 and 4.3 is the following:

> *"On 25 January 1988, CBS Evening News anchor Dan Rather interviews vice-president George H. W. Bush, as part of the coverage of the 1988 presidential election. Before the interview, a video on the Iran–Contra affair was shown to the audience."*

Recall the following annotated segment from Example 4.2:

INTERVIEWER	**(2.1)**	But you made us hypocrites in the face of the world.	**Init-Inform**

The segment conveys a subjective opinion. Assuming that the rest of the dialogue indicates that this is relevant to the subject matter of the interview and that it has not been mentioned before, the following content features are selected for this segment:

<u>On-Topic</u>	**Off-Topic**
Objective	**<u>Subjective</u>**
<u>Accurate</u>	**Inaccurate**
<u>New</u>	**Repeated**

Note that accuracy of the statement could not be checked in this case, so we apply the charitability criterion and judge it as accurate.

Selecting Init-InfoReq Content Features

Example 4.6. Back to the Example 4.1, consider the following question posed a few turns after the quotations in segments **(1.1)** and **(1.2)**:

INTERVIEWER	**(6.1)**	Did you threaten to overrule him?	**Init-InfoReq**

This question is requesting information related to the topic of the interview. It is also neutral (yet sensitive) and reasonable, as it is in the power of the interviewee to provide a reply. Assuming that this is the first time the question is asked, the following content features are selected:

<u>On-Topic</u>	**Off-Topic**
<u>Neutral</u>	**Loaded**
<u>Reasonable</u>	**Unreasonable**
<u>New</u>	**Repeated**

Example 4.7. Now, consider the interview context below:

> *"BBC presenter Jeremy Paxman interviews Conservative MP George Galloway shortly after his parliamentary victory over Labour's Oona King in the UK 2005 General Election."*

The following annotated segment initiates the dialogue:

INTERVIEWER	**(7.1)**	Mr Galloway, are you proud of having got rid of one of the very few black women in Parliament?	**Init-InfoReq**

(*Election Night*, BBC, 2005)

This question is clearly conveying controversial assumptions and is even accusatory. The topic, however, is related to the context of the interview and therefore the following content features are selected:

<u>**On-Topic**</u>	**Off-Topic**
Neutral	<u>**Loaded**</u>
<u>**Reasonable**</u>	**Unreasonable**
<u>**New**</u>	**Repeated**

It must be noted that, although the question is loaded, it is considered reasonable, as it would be possible for the interviewee to provide a satisfactory answer.

Example 4.8. For an instance of unreasonable information request, consider the following context:

> *"In February 2012, BBC Sunday Politics presenter Andrew Neil interviews UK Cabinet Minister Eric Pickles on the Coalition Government's plans for reforms to the National Health Service."*

In the annotated exchange below, as the interviewee notes, it is not in his power to answer the question:

IR	**(8.1)**	Do you deny that three cabinet ministers urged this Conservative Home blog to call for the bill to be junked or emasculated?	**Init-InfoReq**	
IE	**(8.2)**	Er, I have no knowledge of the internal workings of, of Conservative Home"	**Resp-Reject**	@**(8.1)**

(*Sunday Politics*, BBC, 2012)

Therefore, the following content features are selected for segment **(8.1)**:

<u>On-Topic</u>	**Off-Topic**
<u>Neutral</u>	**Loaded**
Reasonable	**<u>Unreasonable</u>**
<u>New</u>	**Repeated**

Selecting Resp-Inform Content Features

Information-giving responsive segments are judged in a way similar to initiating ones. However, in this case the relevance of the topic is judged with respect to the segment to which they respond and not only with respect to the topical context of the interview. The aim is to judge, for instance, whether the information provided by the segment is relevant to the request that motivated it.

Example 4.9. Going back to the interview in Example 4.1, consider the following fragment further into the conversation:

IR	**(9.1)**	Did you threaten to overrule him?	**Init-InfoReq**	
IE	**(9.2)**	I did not overrule Derek Lewis.	**Resp-Inform**	@**(9.1)**

Although the distinction is subtle, the information given in the response is not relevant to the question and the content features below are selected for segment (9.1), assuming that the interviewee has not said this before:

Relevant	**Irrelevant**
Objective	**Subjective**
Accurate	**Inaccurate**
New	**Repeated**
Complete	**Incomplete**

A second difference between the content features of responsive and initiating information giving segments relates to the amount of information provided. Questions usually ask for clearly identifiable pieces of information. *Yes/No*-questions, for instance, can be answered with an affirmative or negative statement (e.g. *"Yes"* or *"No"*), but many times an elaboration is expected. *Wh*-questions ask for one or more objects, individuals, places, and so forth to be identified. Open questions request for positions or opinions on a certain issue. In each case, if annotators are able to determine the amount of information that has been asked for in the segment to which a **Resp-Inform** refers in the annotation, they should be able to decide whether it satisfies the request or not. If it does, then the **Complete** content feature is selected. Otherwise, **Incomplete** is the correct choice. On occasion, the information can be spread across several segments, none of which on its own contains the totality of the information requested. In these cases, the **Incomplete** content feature is selected for all the segments but the last one in the sequence, for which the **Complete** content is chosen.

Example 4.10. Consider the following context:

> *"In February 2011, Channel 4 News presenter Krishnan Guru-Murthy interviews George Osborne, as he attends a G20 meeting of finance ministers in Paris, on the state of the outcomes of the meeting and the state of the British economy."*

In the fragment below, segments **(10.3)** to **(10.5)** have all responsive information giving functions and they are annotated as responding to segment **(1.2)**:

IR	**(10.1)**	So, George Osborne, there you are in Paris with the finest economic minds of the G20.	**Init-Inform**	
	(10.2)	Have you solved the problem of rising food prices?	**Init-InfoReq**	
IE	**(10.3)**	Well, we did talk about the problem of rising food prices and we came up with some of the solutions.	**Resp-Inform**	**@(10.2)**
	(10.4)	Obviously, you can't solve a problem like that overnight, but by giving more information out there about the real cost of things, by trying to promote freer trade, by making sure that some of the poorest producers in the world, in Africa and Asia, get help, financial help to improve their agriculture, what we are trying to do is create more food supply in the world,	**Resp-Inform**	**@(10.2)**
	(10.5)	and that has a real impact on the families in Britain, because, like many other families around the world, we've seen food prices go up.	**Resp-Inform**	**@(10.2)**

(*Channel 4 News*, UK Channel 4, 2011)

Let us see how we annotate each segment bearing in mind the instructions above. Although segment **(10.3)** is relevant to question **(10.2)**, it does not provide all the information requested. For this reason, segment **(10.3)** is annotated with the following content features:

Relevant	**Irrelevant**
Objective	**Subjective**
Accurate	**Inaccurate**
New	**Repeated**
Complete	**Incomplete**

The answer seems to be complete by segment **(10.4)**, where the interviewer admits they have not found a solution, but are working towards it. The content features selected for segment **(10.4)**are as follows:

Relevant	**Irrelevant**
Objective	**Subjective**
Accurate	**Inaccurate**
New	**Repeated**
Complete	**Incomplete**

Segment **(10.5)**, although on a topic related to the context of the interview, is not relevant to the question. The information it conveys has not been requested in segment **(10.2)**. Therefore, the following content features are selected for segment **(10.5)**:

Relevant	**Irrelevant**
Objective	**Subjective**
Accurate	**Inaccurate**
New	**Repeated**
Complete	**Incomplete**

This concludes the description of the annotation approach towards assessment of cooperation. In the next section, we describe a corpus study for evaluating inter-annotator agreement in both annotation stages and discuss the results.

4.3 Evaluation of the Method (Part 1): a corpus annotation study

This section describes an evaluation of the two-stage annotation scheme described above. The method was applied to a corpus of six interview fragments. For the first annotation stage, four annotators received transcripts of the fragments and were asked to segment the turns in each dialogue and to annotate each segment with dialogue act functions and, when applicable, with referent segments. Annotations were automatically aggregated to produce a single segmented and partially annotated version of each dialogue. These were used in the second stage of the study in which seven annotators were asked to select content features. We analysed inter-annotator agreement on the data resulting from each stage to assess the **reliability** of the scheme. Figure 4.7 shows a diagram of the study.

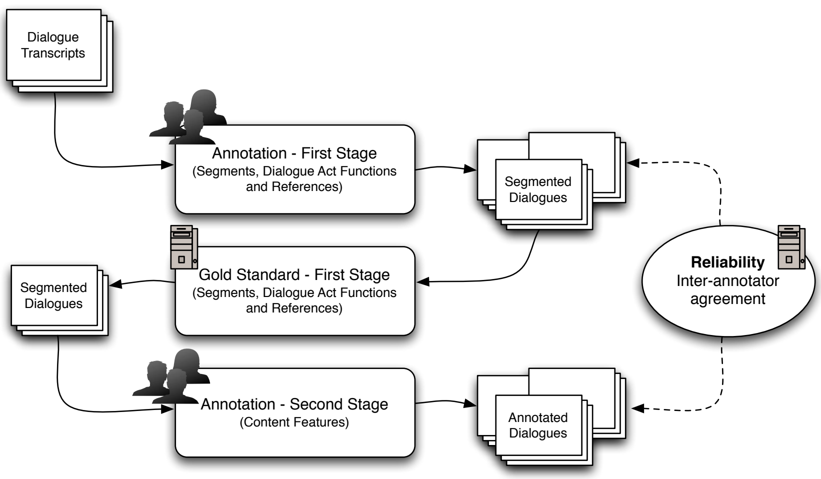


Figure 4.7: Overview of the evaluation approach (Part 1)

4.3.1 Materials

A summary of the materials involved in the annotation study follows. Further detail is given in Appendix A and online at http://mcs.open.ac.uk/nlg/non-cooperation/.

A Corpus of Political Interviews

The study involved the annotation of six interview fragments with a total of 88 turns (3556 words)[11]. The number of turns and words in each fragment is shown in Table 4.1. For further detail and an example of the transcripts, see Appendix A. The entire corpus and annotation data is available at http://mcs.open.ac.uk/nlg/non-cooperation/.

Table 4.1: Political interview fragments in the corpus annotation study

Interview	Turns	Words
1. Brodie and Blair	16	734
2. Green and Miliband	9	526
3. O'Reilly and Hartman	19	360
4. Paxman and Osborne	16	272
5. Pym and Osborne	10	595
6. Shaw and Thatcher	18	1069
Total	88	3556

The fragments were selected from a larger set of 15 interviews collected from publicly available sources (BBC News, CNN, Youtube, etc.; see Appendix A for details). When available, official transcripts from the original sources were used, with minor modifications to reduce the number of functionally empty or split turns (e.g. due to interruptions or overlapped speech). Otherwise, the interviews were transcribed from video or audio clips taken from the sources.

[11] Copyright of all video, audio and transcripts belongs to the respective broadcasting company. Transcripts are reproduced here for the purpose of examination towards an academic degree.

We selected this particular set with the aim of including behaviours at different levels of cooperation on either role. At the same time, we avoided extreme cases in which the exchange broke down or turned into a dialogue of an entirely different type (e.g. confrontation or debate). A second criterion was to ensure coverage of the annotation scheme, with special attention to the dialogue act taxonomy so that all the dimensions in the content feature taxonomy could be tested in the second annotation stage. The third criterion was manageability. In order to reduce the effects of sustained cognitive effort on the quality of annotations, we designed each stage to be completed in one hour. This meant that some of the fragments had to be considerably shortened (e.g. by only including the first few consecutive turns), while keeping their size large enough so that annotators had sufficient context and dialogue history on which to make their judgements[12]. Finally, we limited the interviews in time of broadcast and country of origin. Interviews 3 and 6 were taken from American media (although the politician in Interview 6 is British). The other four fragments involve British interviewers and politicians. All the interviews took place within a span of 14 years, the most dated being from June 1997 (Interview 6) and the most recent from June 2011. These constraints are based on the observation that, as was discussed in Chapter 3, the social conventions underlying a dialogue game change over time and depend highly on cultural background. Considering interviews from a significantly wider time span and from further cultural backgrounds would have interfered with the outcome of the survey study used in the validation of the method.

[12]The fragment of Interview 6 was taken from the end of the interview instead of from the beginning. Care was taken to ensure that the contributions in the fragment did not refer to omitted earlier turns.

Annotation Guidelines

A set of annotation guidelines was prepared for each annotation stage with the definitions and examples presented above. These documents are avaliable online at http://mcs.open.ac.uk/nlg/non-cooperation/. Annotators were given a brief introduction to the instructions and examples before starting their annotations. They read the document in detail and had a chance to clarify any doubts. After these preliminaries, they worked independently on the data.

Annotation Tool

Annotations were carried out using a special-purpose tool, deployed to each annotator containing the annotation data. Among other features, the tool guides annotators through the dialogues in a fixed order and can be configured to operate according to each annotation stage[13].

The main window (Figure 4.8) shows the context of the interview and the transcribed turns with their respective annotations. Clicking on the annotation of a turn opens a window that allows the user to edit the segments and annotations of that turn (Figure 4.9).

The tool also allows collecting information about the background of the annotators (Figure 4.10) and about their familiarity with each dialogue (Figure 4.11 shows an example). The annotation tool user guide with a detailed description of further features and instructions is available at http://mcs.open.ac.uk/nlg/non-cooperation/.

[13] The tool was built based on the CODA *D2MTool* developed by Svetlana Stoyanchev for the CODA Project (Stoyanchev and Piwek, 2010b).

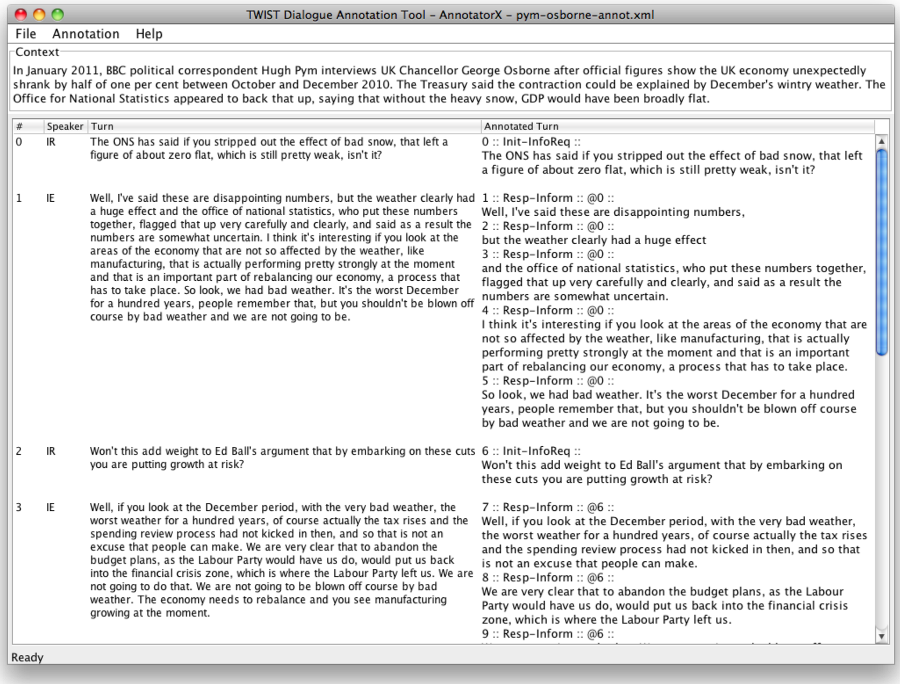


Figure 4.8: Annotation tool: main window

4.3.2 First Annotation Stage: Segmenting Turns and Annotating Dialogue Acts and Referent Segments

In the first stage, annotators received the interview transcripts without any annotations other than the division of turns as spoken by each speaker. Section A.1.1 in the Appendix contains the transcript of Interview 1. The entire corpus is available at `http://mcs.open.ac.uk/nlg/non-cooperation/`.

Annotators

Four annotators were involved in the first stage: the author, one of his supervisors and two native English-speaking researchers with previous experience in dialogue annotation and some familiarity with the topic of this thesis. Before annotating each dialogue, annotators were asked to provide information about their familiarity with the interview, its context and the dialogue participants. A summary of their answers is shown in Table 4.2.

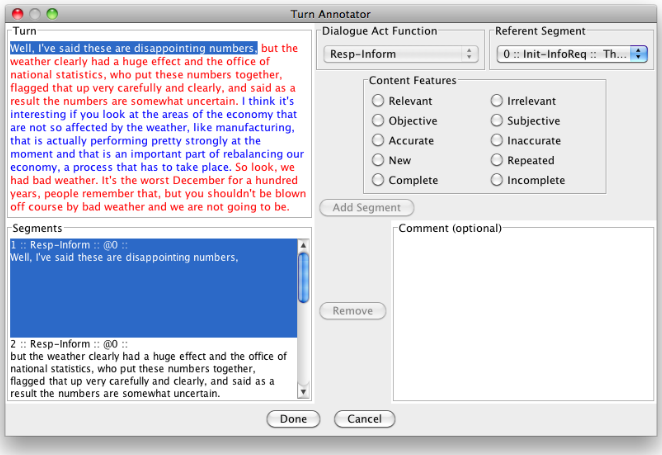


Figure 4.9: Annotation tool: turn annotator window

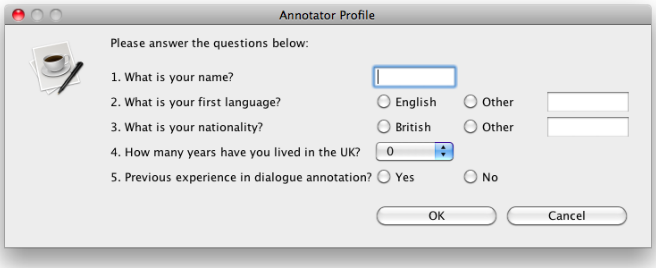


Figure 4.10: Annotator profile

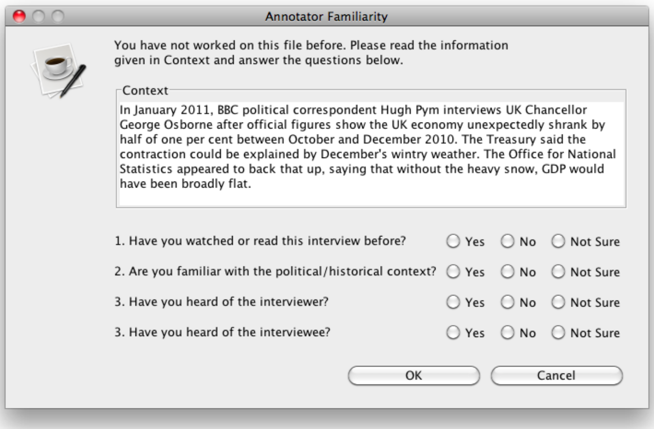


Figure 4.11: Annotator familiarity with respect to an interview

Table 4.2: Familiarity of the annotators with the interviews, their contexts and the dialogue participants (first annotation stage)

Interview	Context			Dialogue			Interviewer			Interviewee		
	Y	N	NS	Y	N	NS	Y	N	NS	Y	N	NS
1. Brodie and Blair	4	0	0	2	1	1	2	2	0	4	0	0
2. Green and Miliband	2	2	0	1	2	1	2	2	0	4	0	0
3. O'Reilly and Hartman	1	2	1	1	3	0	2	2	0	1	3	0
4. Paxman and Osborne	2	2	0	2	2	0	4	0	0	4	0	0
5. Pym and Osborne	4	0	0	2	2	0	2	2	0	4	0	0
6. Shaw and Thatcher	4	0	0	1	3	0	2	2	0	4	0	0

Legend. Y: *'yes'*, **N**: *'no'*, **NS**: *'not sure'*.
The figures represent the number of annotators that chose each option.

Results

Inter-annotator agreement was measured on segmentation and on the assignment of dialogue act functions and referent segments. For segmentation, we used Krippendorff's α_U, a coefficient for assessing the reliability of unitising continuous data (Krippendorff, 1995). Agreement on the assignment of dialogue act functions and referent segments was measured using Krippendorff's α (Krippendorff, 2004) and multi-rater versions of Cohen's κ (Cohen, 1960; Davies and Fleiss, 1982) and Scott's π (Scott, 1955; Fleiss, 1971).

Krippendorff's α is a family of reliability coefficients (Krippendorff, 2003, Chapter 11) defined in terms of the ratio between the disagreement observed among the coders and the disagreement expected by chance:

$$\alpha = 1 - \frac{D_o}{D_e}$$

where D_o and D_e are, respectively, the observed and expected disagreements. A value of $\alpha = 1$ results when $D_o = 0$, that is, when there is perfect agreement among the annotators. A value of $\alpha = 0$ results when $D_o = D_e$, meaning that annotators agreed as expected by chance, thus rendering the annotated data unreliable. Negative values of α mean that annotators disagreed more than is predicted by chance what usually indicates systematic discrepancies in their interpretation of the annotation criteria.

Observed and expected disagreements are calculated based on differences between annotation categories. These differences are given by a distance function $\mathbf{d}$ that can be defined according to the nature of the coding scheme, provided that the distance between identical categories is always zero ($\mathbf{d}_{aa} = 0$), that distances are symmetric ($\mathbf{d}_{ab} = \mathbf{d}_{ba}$) and that the triangular inequality holds ($\mathbf{d}_{ab}+\mathbf{d}_{bc} \geq \mathbf{d}_{ac}$). This spawns a whole family of rather versatile coefficients, with properties that make them suitable for assessing the reliability of our annotation scheme. Krippendorff's α supports, for instance, multiple coders, incomplete data (missing annotations), different kinds of annotations categories (nominal, ordinal, interval, ratio), weighed distance metrics, small samples and unitisation (Krippendorff, 2004).

In addition to Krippendorff's α, we report reliability of the annotation of dialogue act functions using multi-rater versions of **Cohen's** κ (Cohen, 1960; Davies and Fleiss, 1982) and **Scott's** π (Scott, 1955; Fleiss, 1971) – called K by Siegel and Castellan (1988) . This is because these measures are often found in the literature when discussing the results of dialogue annotation exercises. The general form for both coefficients is:

$$\pi, \kappa = \frac{A_o - A_e}{1 - A_e}$$

where A_o and A_e are, respectively, the observed – or average – agreement and the agreement expected by chance. The observed agreement A_o is the same for both coefficients and equal to the ratio between the number of instances in which any two annotators agreed in the classification of an item and the total number of pairs of annotations of each item. The latter is the number of possible pairs of annotators, times the number of items. The former is the number of possible pairs of annotators that classified an item under a label, summed over all items and all labels (Artstein and Poesio, 2008, p. 563):

$$A_o = \frac{1}{\mathbf{i}\binom{\mathbf{c}}{2}} \sum_{i \in I} \sum_{k \in K} \binom{\mathbf{n}_{ik}}{2} = \frac{1}{\mathbf{ic}(\mathbf{c}-1)} \sum_{i \in I} \sum_{k \in K} \mathbf{n}_{ik}(\mathbf{n}_{ik}-1)$$

where $\mathbf{i}$ is the size of the set of items I, $\mathbf{c}$ is the number of annotators, K is the set of labels and $\mathbf{n}_{ik}$ is the number of annotators that classified item i with label k.

The expected agreement A_e is the mean of the probability of any two coders agreeing on assigning a label to any items, summed over all labels. As before, the total number of pairs of annotators is $\binom{\mathbf{c}}{2}$, where $\mathbf{c}$ is the number of annotators. If $\mathrm{P}(k|c)$ and $\mathrm{P}(k|d)$ denote, respectively, the probability of coder c and d assigning label k to an item, then the probability of both coders assigning k to the same item is $\mathrm{P}(k|c) \cdot \mathrm{P}(k|d)$. The expected agreement is then[14]:

$$A_e = \sum_{k \in K} \frac{1}{2\binom{\mathbf{c}}{2}} \sum_{c \in C} \sum_{\substack{d \in C \\ c \neq d}} \mathrm{P}(k|c) \cdot \mathrm{P}(k|d) = \sum_{k \in K} \frac{1}{\mathbf{c}(\mathbf{c}-1)} \sum_{c \in C} \sum_{\substack{d \in C \\ c \neq d}} \mathrm{P}(k|c) \cdot \mathrm{P}(k|d)$$

where K is the set of labels and $\mathbf{c}$ is the size of the set of coders C.

The coefficients differ in the way the probabilities $\mathrm{P}(k|c)$ and $\mathrm{P}(k|d)$ are calculated. In the absence of a probability distribution for the labels in the annotation scheme, the probabilities are estimated based on the annotation data. For Scott's π – or Siegel and Castellan's K –, it is assumed that the

[14]The formula is based on Artstein and Poesio (2008, p. 560). The mean is over twice the number of coder pairs, $2\binom{\mathbf{c}}{2}$, as the summation is over all pairs of different coders twice.

probability of a coder choosing a given label by chance is the same for all coders and equal to the proportion of items assigned to label k over the total number of assignments. That is, $\mathrm{P}(k|c) = \mathrm{P}(k|d) = \hat{\mathrm{P}}(k) = {}^{\mathbf{n}_k}/_{\mathbf{ic}}$, where $\mathbf{n}_k$ is the number of items classified under label k, $\mathbf{i}$ is the number of items and $\mathbf{c}$ is the number of coders. In Cohen's κ, on the other hand, a different probability distribution is estimated for each coder based on their individual annotations. That is, $\mathrm{P}(k|c) = \hat{P}(k|c) = {}^{\mathbf{n}_{ck}}/_{\mathbf{i}}$, where $\mathbf{n}_{ck}$ is the number of items classified under label k by coder c and $\mathbf{i}$ is the total number of items[15].

From the definitions of A_o and A_e it follows that κ and π can only be used when all items have been annotated by all coders[16]. In this case, it can be shown that the observed disagreement in Krippendorff's α is equal to the complement of observed agreement, $D_o = 1 - A_o$, with **binary difference** as the distance function $\mathbf{d}$:

$$\mathbf{d}_{ab} = \begin{cases} 0 & \text{if } a = b \\ 1 & \text{if } a \neq b \end{cases}$$

Moreover, with this distance function, expected disagreement in Krippendorff's α is closely equivalent to the complement of expected agreement for π (or K) so these two coefficients take similar values as the number of items or annotators grow larger (Artstein and Poesio, 2008, p. 567). When possible, however, we report agreement values for the three coefficients, as κ and π/K are often found in related work and α can be applied consistently across the corpus and for the entire coding scheme.

[15] For further details and a discussion on the properties of these coefficients, we refer the reader to Artstein and Poesio's article (Artstein and Poesio, 2008, pp. 559–562).

[16] Note that A_o is defined as an average over the number of items and pairs of coders which is only accurate if all annotators annotated every item. For coded data with missing annotations, we measure agreement using only Kripendorff's α.

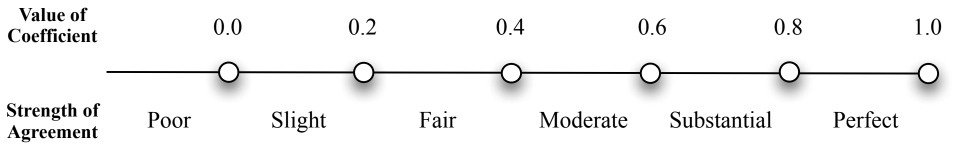


Figure 4.12: Values of κ-like coefficients and strength of agreement according to Landis and Koch (1977) adapted from Artstein and Poesio (2008, p. 576).

There is no consensus regarding how high a value of any of the coefficients renders annotated data reliable for further use. Artstein and Poesio (2008, pp. 576–577) discuss this and state that no clear threshold has been established in the field of computational linguistics since reliability studies were first introduced to the community by Carletta et al. (1997). The authors refer to medical literature where values for κ-like coefficients above 0.4 are considered adequate and point to similar conventions in the field of language studies (see Figure 4.12). They warn, however, that in computational linguistics, higher thresholds are usually adopted. They quote Carletta et al. (1997) who follow Krippendorff (1980, p. 147) in that only annotated data with values of α above 0.8 should be considered reliable, with values between 0.67 and 0.8 allowing for "tentative conclusions". In their conclusions, Artstein and Poesio (2008) settle for 0.8 as a good value if a threshold is required. They are uncertain, however, that this single value could be used as a cutoff point for all purposes and, in agreement with Di Eugenio and Glass (2004) and others recommend that researchers "report in detail on the methodology that was followed in collecting the reliability data (number of coders, whether they coded independently, whether they relied exclusively on an annotation manual), whether agreement was statistically significant, and provide a confusion matrix or agreement table so that readers can find out whether overall figures of agreement hide disagreements on less common

Table 4.3: Number of segments identified by each annotator

Interview	Annot.1	Annot.2	Annot.3	Annot.4	Total
1. Brodie and Blair	21	24	20	21	86
2. Green and Miliband	16	25	13	15	69
3. O'Reilly and Hartman	29	30	27	30	116
4. Paxman and Osborne	20	17	15	13	65
5. Pym and Osborne	13	20	11	16	60
6. Shaw and Thatcher	25	28	20	37	110
Total	124	144	106	132	506

categories." (Artstein and Poesio, 2008, p. 591). This is what we present in the rest of this section. All annotation data is available from the author in text and digital format.

Inter-Annotator Agreement for Segmentation. Table 4.3 summarises the number of segments identified by the annotators in each dialogue.

Krippendorff (1995; 2004) calls **unitising** the process of identifying the units of annotation in continuous data. This is similar to what we call segmentation, except that in our case the data is continuous only within a turn. Furthermore, Krippendorff's units exhaustively cover the continuum of data, while segments refer only to functionally "interesting" sections of speech, with uninteresting sections left unmarked. Nevertheless, Krippendorff's definition allows for certain units to be discarded in posterior annotations. Discarded units are analogous to gaps between segments in our approach, except that two or more consecutive discarded units retain their respective boundaries, while in our annotations they are regarded as a single gap.

Disagreement is computed by pairing the units identified by each annotator with those identified by all other annotators, and measuring how much paired units that overlap differ from each other. Artstein and Poesio concisely describe the metric used for measuring this difference as follows (Artstein and Poesio, 2008, p. 582):

> "If a unit identified by one coder overlaps a unit identified by the other coder, the amount of disagreement is the square of the lengths of the non-overlapping segments [...]; if a unit identified by one coder does not overlap any unit of interest identified by the other coder, the amount of disagreement is the square of the length of the whole unit."

Krippendorff (1995) provides a geometric interpretation of the metric and of observed and expected disagreement. We implemented these as proposed by the author, measuring starting positions and lengths of segments in number of words, with a modification to consider dialogue data that is divided in turns, rather than as a single continuum. This change is of relevance, for instance, when computing expected disagreement as this requires considering all possible units and gaps that could overlap with a given unit. In a single continuum of data the boundaries of such units and gaps can appear anywhere, while in our data they are limited within a single turn. Thus, for each dialogue in the corpus, observed and expected disagreement are the sum of the observed and expected disagreement in each turn (t):

$$\alpha_U = 1 - \frac{\sum_t D_{ot}}{\sum_t D_{et}}$$

Similarly, overall agreement for the entire corpus can be obtained by adding the respective disagreements of each dialogue. The results are presented in Table 4.4.

In general, agreement for segmentation is high ("substantial", in terms of the scale in Figure 4.12). Consistent with intuition, disagreement is greater in dialogues with longer turns.

Artstein and Poesio (2008) argue that α_U is not applicable to all tasks

Table 4.4: Inter-annotator agreement for segmentation (Krippendorff's α_U)

Interview	α_U	D_o	D_e
1. Brodie and Blair	0.802	3.217	16.251
2. Green and Miliband	0.618	3.276	8.565
3. O'Reilly and Hartman	0.773	4.138	18.219
4. Paxman and Osborne	0.92	0.993	12.468
5. Pym and Osborne	0.672	4.0	12.184
6. Shaw and Thatcher	0.653	7.951	22.890
Overall	0.74	23.574	90.577

in computational linguistics. Among the reasons for this they mention the method's assumption that units identified by a single coder do not overlap. They note, however, that when this assumption holds and when parts of the data can be left unannotated, testing the reliability of segmentation can be beneficial (Artstein and Poesio, 2008, p. 583). They also note that, as far as they know, this had not been tested in computational linguistics at the time of publication of their article[17]. Our annotation scheme meets these assumptions. This made testing agreement attractive, especially as we introduced some uncertainty by including topicality among the criteria for choosing segment boundaries.

Inter-Annotator Agreement for Dialogue Act Functions. Table 4.5 summarises the number of occurrences of each dialogue act function over the 506 segments identified in total by the four coders (frequencies are given in parentheses).

These frequencies are consistent with the assumptions in Chapter 3. In general, political interviews are predominantly made up of questions and

[17]The authors refer to several studies on the reliability of segmentation using Siegel and Castellan's K and the presence or absence of boundaries as possible classifications of the space between words, as reported for instance by Carletta et al. (1997). They note, however, that using such coefficients for analysing segmentation can lead to values of agreement that are artificially low (Artstein and Poesio, 2008, pp. 580–582). This is because the boundary/not boundary distinction makes no consideration for cases in which two annotators agreed in most of a segment but disagreed on the exact boundaries.

Table 4.5: Number of occurrences and frequency of dialogue act functions in the annotated corpus

Interview	Init-Inform	Init-InfoReq	Resp-Inform	Resp-Accept	Resp-Reject	Total
1. Brodie and Blair	8 (0.093)	33 (0.384)	43 (0.500)	1 (0.012)	1 (0.012)	86
2. Green and Miliband	24 (0.348)	22 (0.319)	19 (0.275)	1 (0.014)	3 (0.043)	69
3. O'Reilly and Hartman	27 (0.233)	9 (0.078)	28 (0.241)	14 (0.121)	38 (0.328)	116
4. Paxman and Osborne	12 (0.185)	16 (0.246)	16 (0.246)	8 (0.123)	13 (0.200)	65
5. Pym and Osborne	6 (0.100)	32 (0.333)	32 (0.533)	1 (0.017)	1 (0.017)	60
6. Shaw and Thatcher	15 (0.136)	47 (0.291)	47 (0.427)	9 (0.082)	7 (0.064)	110
Total	92 (0.182)	132 (0.261)	185 (0.366)	34 (0.067)	63 (0.125)	506

replies, with almost two thirds of the dialogue acts falling in either of these categories. Replies are slightly longer than questions extending over several segments, and about one fifth of the contributions are initiating statements – e.g. preambles to questions or comments by either speaker that are not reactions to previous questions or replies. Explicit rejections are twice as frequent as explicit acceptances, supporting the assumption that acceptances are usually left implicit. These frequencies vary dramatically for some of the dialogues. A rather extreme example is Dialogue 3, in which initiating statements are three times more abundant than questions, the number of explicit acceptances doubles the average for the corpus, and almost a third of the segments are explicit rejections.

For computing the values of the agreement coefficients presented in the remainder of the section, we based our implementation on a Python script developed by Tom Lippincott[18]. A later version of this script has recently been included in the *Natural Language Toolkit* (NLTK) suite of libraries for natural language processing in Python (Bird et al., 2009)[19].

As noted above, annotators segmented the turns and selected dialogue

[18]The original Python script was taken from *The Alpha resources page*, a companion page to the article by Artstein and Poesio (2008): `http://cswww.essex.ac.uk/Research/nle/arrau/alpha.html` (accessed: December 2012). The modified version of the script is available from the author of this thesis.

[19]The NLTK version of Lippincott's script is available at `http://nltk.org/_modules/nltk/metrics/agreement.html` (accessed: December 2012).

Table 4.6: Inter-annotator agreement for dialogue act functions

Label	α	D_o	D_e
Init-Inform	0.409	0.040	0.068
Init-InfoReq	0.893	0.009	0.089
Resp-Inform	0.645	0.038	0.107
Resp-Accept	0.606	0.011	0.029
Resp-Reject	0.635	0.018	0.050
Overall	0.657	0.059	0.171

Table 4.7: Pairwise confusion matrix of dialogue act function annotations

	Init-Inform	Init-InfoReq	Resp-Inform	Resp-Accept	Resp-Reject
Init-Inform	40	9	40	5	10
Init-InfoReq		145	1	3	2
Resp-Inform			68	6	13
Resp-Accept				15	4
Resp-Reject					36

act functions for these segments independently and in the same annotation step. This means that the units for annotation identified by one coder can differ from those identified by another coder. Figure 4.13(a) illustrates this: the three annotators coincided in identifying the initial segment, but differed in the remainder of the turn. These differences make it possible to analyse the reliability of the original annotation data only in terms of Krippendorff's α, which supports missing annotations for some of the items. The value of this coefficient for each label (i.e. regarding the rest of the categories as **'Other'**) and for entire dialogue act taxonomy is given in Table 4.6.

Agreement ranges from "moderate" to "perfect" (see the scale in Figure 4.12), with overall agreement being "substantial". Table 4.7 shows a **condensed pairwise confusion matrix**[20].

Recall that in total the four annotators identified 506 segments. Of these, 138 were identified by at most one annotator and are not part of the con-

[20]The figures in each cell represent the number of instances in which two coders assigned the labels in the cell's row and column to the same segment. This considers each pair of annotators only once for each segment (thus the upper triangular shape of the matrix).

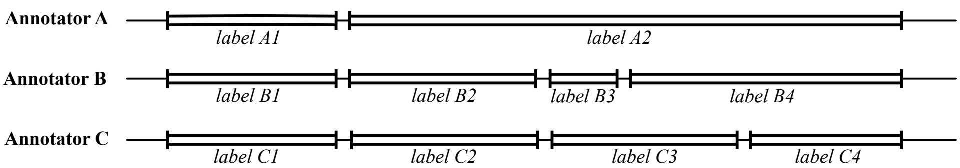
fusion matrix. The remaining 368 segments were annotated by two or more coders. This resulted in 397 pairs of annotations of the same element by two different coders (i.e. the sum of all the elements in the matrix), 304 of which are on the diagonal of the matrix – meaning that both annotators in the pair coincided in their classification of the segment. The highest number of mismatches (40) happened between initiating and responsive information-giving dialogue acts (**Init-Inform** and **Resp-Inform**, respectively). The function of both dialogue acts is that of providing information, with the difference that the latter is doing so in response to a previous act while the former is initiating an exchange or providing context for an upcoming question. These results could be improved by further clarification in the guidelines, for instance, emphasising that attention should first focus on whether an act is responsive or initiating and an explicit mention on how to deal with information-giving dialogue acts that are given in response to a question but that do not constitute a relevant answer. The next highest number of mismatches happened between **Resp-Reject** and **Resp-Inform** or **Init-Inform**. This might call for further clarification in the guidelines as to what should be regarded as an explicit rejection.

As stated, overall agreement was acceptably high. It is unclear, however, how these results compare to related work in the annotation of dialogue acts, given the differences in the identification of segment boundaries by individual coders we discussed above. Although Krippendorff's α deals with incomplete annotations, in this case the missing annotations are not caused by coders omitting the classification of some items, but rather by these items being different across annotators. Carletta et al. (1997, p. 26) note that only data where segmentation is robust should be analysed further. Their solution is to measure reliability of dialogue act classification only in segments whose boundaries were identified by all coders. This is along the lines of what we do below, but the same arguments that support the use of Krippendorff's α_U for analysing segmentation agreement as opposed to Siegel and Castellan's K apply here. Eliminating segments just because not all annotators agreed exactly on their boundaries is too harsh and reduces the dataset considerably.

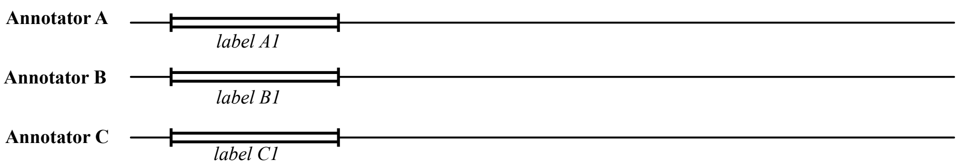
A straightforward way of avoiding these problems is to carry out the segmentation and the classification of dialogue act functions in two separate annotation stages. Time constraints prevented us from repeating the annotation exercise, but this modification to the scheme is recommended for future studies. The reason why we did not do it in the first place is that segments are defined based on their function, which is precisely what dialogue act functions encode. This made it natural to combine both annotation tasks in one step. However, nothing would prevent us from adding an intermediate stage in which the annotation of dialogue act functions from the first stage are discarded, segmentation from all annotators is aggregated into a single set as explained in the next section, and segmented dialogues are given to a (potentially new) set of annotators who will then select the dialogue act functions and referent segments.

In order to apply the other two coefficients to the annotated dialogues, we considered two derived sets of annotations in which all segments received four judgements. In the first set we followed a criterion called **strict segment matching**: a segment and its annotation are retained in the derived set if *all* coders selected that segment in their annotation. This is illustrated (for three coders) by the initial segment in Figure 4.13(a), resulting in the subset of annotations in Figure 4.13(b).

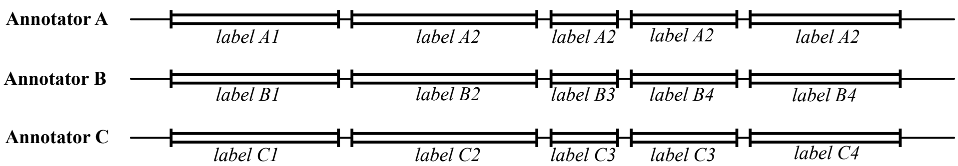
Of the 506 segments identified in total by the four annotators, 156 be-



(a) Original annotation data



(b) Strict segment matching



(c) Lenient segment matching

Figure 4.13: A turn coded by three annotators and the approach for measuring reliability of dialogue act functions: (a) original annotated segments; (b) subset of segments matching exactly across all annotators; (c) lenient matching considering subsegments.

long in this subset – that is, 39 distinct segments, each one with four annotations. Table 4.8 shows the frequencies as before but for the annotations in this subset. With respect to the original set of annotations, the most salient differences in the frequencies are in questions and replies (labels **Init-InfoReq** and **Resp-Inform**, respectively). While replies went down from over a third of the items to just above 15%, questions doubled in frequency to over half the total number of segments. This is explained by noting that question turns tend to be shorter than reply turns, thus making annotators agree more often in the identification of these segments. On the other hand, longer reply turns increase the chances that annotators select different boundaries for the segments within, which are thus excluded from this

Table 4.8: Number of occurrences and frequency of dialogue act functions in strictly matching annotated segments

Interview	Init-Inform	Init-InfoReq	Resp-Inform	Resp-Accept	Resp-Reject	Total
1. Brodie and Blair	0 (0.00)	28 (0.875)	4 (0.125)	0 (0.000)	0 (0.000)	32
2. Green and Miliband	1 (0.063)	12 (0.75)	3 (0.188)	0 (0.000)	0 (0.000)	16
3. O'Reilly and Hartman	10 (0.227)	8 (0.182)	1 (0.023)	9 (0.205)	16 (0.364)	44
4. Paxman and Osborne	4 (0.500)	4 (0.500)	0 (0.000)	0 (0.000)	0 (0.000)	8
5. Pym and Osborne	5 (0.179)	4 (0.143)	12 (0.429)	1 (0.036)	6 (0.214)	28
6. Shaw and Thatcher	1 (0.036)	23 (0.821)	4 (0.143)	0 (0.000)	0 (0.000)	28
Total	21 (0.135)	79 (0.506)	24 (0.154)	10 (0.064)	22 (0.141)	156

Table 4.9: Inter-annotator agreement for dialogue act functions in the set of strictly matching annotated segments

Label	α (D_o, D_e)	A_o	κ (A_e)	$\pi \vert K$ (A_e)
Init-Inform	0.471 (0.124, 0.234)	0.876	0.470 (0.766)	0.468 (0.767)
Init-InfoReq	0.924 (0.038, 0.503)	0.962	0.923 (0.500)	0.923 (0.500)
Resp-Inform	0.674 (0.085, 0.262)	0.915	0.673 (0.739)	0.672 (0.740)
Resp-Accept	0.575 (0.051, 0.121)	0.949	0.573 (0.880)	0.573 (0.880)
Resp-Reject	0.790 (0.051, 0.244)	0.949	0.788 (0.758)	0.788 (0.758)
Overall	0.743 (0.175, 0.682)	0.825	0.742 (0.321)	0.741 (0.322)

subset. Table 4.9 presents the values for Krippendorff's α, observed and expected disagreement, observed (or average) agreement A_o, and multi-rater versions of Cohen's κ and Scott's π (or Siegel and Castellan's K) with their respective expected agreements A_e – recall that observed agreement is the same for both coefficients and as given under A_o.

The table shows that the three coefficients take values that are very close to each other for all labels. With the exception of **Resp-Accept**, all the values for Krippendorff's α are slightly higher here than in the original set of segments. Agreement ranges from "moderate" to "perfect", with overall agreement being "substantial". The overall score of $K = 0.741$ is lower than those reported for similar dialogue act taxonomies. Carletta et al. (1997) reported an overall agreement of $K = 0.83$ for the MapTask coding scheme. A one-dimensional version of the DAMSL scheme for the Switchboard domain by Jurafsky et al. (1997) resulted in overall agreement of $K = 0.8$. Stoyanchev and Piwek (2010b) obtained overall agreement of $K = 0.82$ for

Table 4.10: Pairwise confusion matrix of dialogue act function annotations for strictly matching segments

	Init-Inform	**Init-InfoReq**	**Resp-Inform**	**Resp-Accept**	**Resp-Reject**
Init-Inform	17	6	15	3	5
Init-InfoReq		114	0	3	0
Resp-Inform			26	2	3
Resp-Accept				9	4
Resp-Reject					27

their coding scheme for expository dialogues. Bunt (2009) reported on near-perfect agreement ($K = 0.98$ in average for most categories) on a study of the LIRICS taxonomy – a predecessor of the ISO standard for dialogue act annotation proposed later by Bunt et al. (2010; 2012). Although our values are lower, especially for initiating information-giving acts and for explicit acceptances, they are encouragingly high considering that we had a single run of annotations, that there was no training for non-expert annotators and that the dialogues were open-domain.

Table 4.10 shows the condensed pairwise confusion matrix for this set of annotations. There are six pairs of judgements for each of the 39 segments what amounts to 234 pairs of judgements. Of these, 193 lie on the diagonal of the matrix, meaning that two annotators coincided in the choice of label for a segment. The remaining 41 pairs are mismatches. The majority of these (15) happened between the labels **Init-Inform** and **Resp-Inform**. In fact, 29 of the 41 mismatches involved **Init-Inform** which explains the low agreement score for this label ($K = 0.468$).

To compute κ and π/K on a larger portion of the annotated corpus, we considered a second derived set of segments by applying a criterion called **lenient segment matching**: for each annotator, an opening or closing boundary is introduced in the derived set of segments if *at least one* of the coders included that boundary in their annotations, provided that the resulting segment is equal to or contained by the segments identified by all other annotators[21]. The intention behind this criterion is to allow for portions of a turn that have been annotated by all coders but that were partitioned in different ways during segmentation to be included in the analysis of agreement for dialogue act labels using κ and π/K. Figure 4.13(a) illustrates this. All the annotators agreed on the first segment of the turn, but thereafter annotator A identified one large segment, while annotators B and C identified three (agreeing on the boundaries of the first of these segments, while disagreeing on those of the other two). Thus, according to the criterion, larger segments are partitioned as illustrated in Figure 4.13(c) and agreement is assessed by comparing the respective annotation labels.

Applying the lenient criterion to the annotated dataset resulted in 696 judgements, that is 174 segments or subsegments, each one with four judgements. Table 4.11 shows the frequencies of the labels for each dialogue and for the entire corpus. These frequencies are similar to those in the original set of annotations. There are more segments classified as **Init-Inform** (148, from 92 in the original set) and as **Resp-Inform** (307, from 185 in the original set). This is because initiating statements and responses to questions are generally longer than questions, acceptances and rejections which can lead to higher disagreement in the exact boundaries of the segments. This is the same reason why most of these segments were filtered out by the strict segmentation criterion (only 21 segments labelled as **Init-Inform** and 24 labelled as **Resp-Inform** were part of the resulting set).

Table 4.12 shows the values of the three agreement coefficients for this set

[21] The containment condition is required to make sure that all the segments in the resulting set are annotated by all coders. When there is a gap because one of the annotators left a portion of a turn unsegmented, then that portion is omitted from the derived set according to the criterion, as the definition of κ and π/K require that items are annotated by all coders.

Table 4.11: Number of occurrences and frequency of dialogue act functions in leniently matching annotated segments

Interview	Init-Inform	Init-InfoReq	Resp-Inform	Resp-Accept	Resp-Reject	Total
1. Brodie and Blair	34 (0.293)	14 (0.121)	65 (0.56)	1 (0.009)	2 (0.017)	116
2. Green and Miliband	23 (0.25)	9 (0.098)	55 (0.598)	1 (0.011)	4 (0.043)	92
3. O'Reilly and Hartman	9 (0.063)	34 (0.236)	51 (0.354)	12 (0.083)	38 (0.264)	144
4. Paxman and Osborne	28 (0.28)	35 (0.35)	28 (0.28)	1 (0.01)	8 (0.08)	100
5. Pym and Osborne	21 (0.328)	12 (0.188)	16 (0.25)	1 (0.016)	14 (0.219)	64
6. Shaw and Thatcher	33 (0.183)	30 (0.167)	92 (0.511)	14 (0.078)	11 (0.061)	180
Total	148 (0.213)	134 (0.193)	307 (0.441)	30 (0.043)	77 (0.111)	696

Table 4.12: Inter-annotator agreement for dialogue act functions in the set of leniently matching annotated segments

Label	α (D_o, D_e)	A_o	κ (A_e)	$\pi\|K$ (A_e)
Init-Inform	0.317 (0.213, 0.311)	0.787	0.333 (0.681)	0.316 (0.689)
Init-InfoReq	0.863 (0.046, 0.335)	0.954	0.863 (0.665)	0.863 (0.665)
Resp-Inform	0.494 (0.250, 0.494)	0.750	0.498 (0.502)	0.493 (0.507)
Resp-Accept	0.258 (0.061, 0.083)	0.939	0.258 (0.917)	0.257 (0.918)
Resp-Reject	0.480 (0.102, 0.197)	0.898	0.482 (0.802)	0.479 (0.803)
Overall	0.527 (0.336, 0.710)	0.664	0.530 (0.284)	0.526 (0.291)

of segments. Except for **Init-InfoReq**, agreement values for all the labels are significantly lower than those we obtained for the original set of annotations and for the strict matching.

In Table 4.13 we give the condensed pairwise confusion matrix for this set of judgements. The six pairs of judgements for each of the 174 segments yield a total of 1044 pairs, 693 of which are matching pairs. The remaining 351 pairs are mismatches. It emerges that confusion is high between **Init-Inform** and **Resp-Inform**, as before, but also between **Resp-Inform** and **Resp-Accept** or **Resp-Reject**. The confusion between initiating and responsive information-giving acts is worsened here by the length of

Table 4.13: Pairwise confusion matrix of dialogue act function annotations for leniently matching segments

	Init-Inform	Init-InfoReq	Resp-Inform	Resp-Accept	Resp-Reject
Init-Inform	90	34	153	7	28
Init-InfoReq		198	1	5	8
Resp-Inform			330	44	63
Resp-Accept				13	8
Resp-Reject					62

the segments which, as said, results in higher disagreement on the segmentation and thus in more subsegments with mismatching judgements. The high number of mismatches between replies and explicit acceptances and rejections results from annotators that identified longer segments considering acceptances and rejections implicit – and therefore part of an information-giving act – while annotators favouring shorter segments would classify the first one as an explicit acceptance or rejection followed by one or more responsive information-giving acts.

Inter-Annotator Agreement for Referent Segments. In Section 4.2.2 and in Appendix A, we used sequential numbering for identifying previous segments in the selection of referents for responsive dialogue acts. However, as segmentation differs from coder to coder, sequential numbering is only consistent within the annotations of a single coder. For illustration, consider the fragments annotated by different coders shown in Figure 4.14. While coder A identified one segment in Turn 1, coder B identified three. This means that the numberings of subsequent segments are different and, for instance, although both annotators chose the question in Turn 2 as referent of the last segment in Turn 3, the actual values are not the same. For comparing the choice of referents across annotators, we identify segments using the index of the turn they belong in (t) and the starting (s) and ending (e) boundaries (in number of characters from the beginning of the turn), using colons as separators: ($t : s : e$). In the example, the segment numbered as (2) by coder A and as (4) by coder B is identified as ($2 : 0 : 28$). This guarantees that we can compare referent segment chosen by different annotators independently of how many segments they identified earlier in the dialogue.

Still, comparing referent segments is not a trivial task. The binary dif-

Turn	Spkr.	Annotated Speech
0	IR	**(0)[Init-InfoReq** Is Osama Bin Laden your prime suspect?]
1	IE	**(1)[Resp-Inform@(0)** He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.]
2	IR	**(2)[Init-InfoReq** Him alone or anybody else?]
3	IE	**(3)[Resp-Inform@(2)** Well, when we assemble the evidence finally, we will present it to people.] **(4)[Resp-Inform@(2)** But as we have said he is the prime suspect.]

(a) Fragment annotated by coder A

Turn	Spkr.	Annotated Speech
0	IR	**(0)[Init-InfoReq** Is Osama Bin Laden your prime suspect?]
1	IE	**(1)[Resp-Inform@(0)** He is the prime suspect.] **(2)[Init-Inform** We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam.] **(3)[Resp-Inform@(0)** And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.]
2	IR	**(4)[Init-InfoReq** Him alone or anybody else?]
3	IE	**(5)[Init-Inform** Well, when we assemble the evidence finally, we will present it to people.] **(6)[Resp-Inform@(4)** But as we have said he is the prime suspect.]

(b) Fragment annotated by coder B

Figure 4.14: Example of a fragment annotated by two coders in which segmentation differs

ference distance function used above for dialogue act functions can be too strict. It requires exact matches between segment boundaries, so two referent segments overlapping almost perfectly except for a word, for instance, would constitute a mismatch[22]. Therefore, in addition to the binary difference distance function, we report here the results for an alternative, more relaxed binary metric which we called **overlapping**. The overlapping distance function between two referent segments in the same dialogue is 0 if segments overlap and 1 otherwise[23].

[22]These problems would be solved, at least in part, by separating segmentation and dialogue act function and referent segments annotation in two stages as suggested above. In that way, targets could be regarded as labels and binary difference could be used safely. However, the problem of estimating expected agreement more accurately by considering the restrictions in the options available to annotators at each point in a dialogue would remain.

[23]We tried a third metric that considered the proportion of overlap between the segments

Table 4.14: Inter-annotator agreement for referent segment annotations

Segments	α^e (D_o, D_e)	α^o (D_o, D_e)	A_o	κ (A_e)	$\pi\|K$ (A_e)
Original	0.732 (0.038, 0.141)	0.699 (0.007, 0.024)	–	–	–
Strict	0.459 (0.519, 0.959)	0.780 (0.074, 0.337)	0.481	0.452 (0.068)	0.444 (0.068)
Lenient	0.646 (0.337, 0.952)	0.788 (0.074, 0.210)	0.663	0.646 (0.049)	0.644 (0.053)

Table 4.14 summarises agreement values for α with the exact (α^e) and overlap (α^o) binary metrics, κ and π/K for the original set of segments and also, as before, for strict and lenient matching sets of segments. Of the 506 segments identified in total by the four coders, 282 were annotated as responsive and assigned referent segments (recall Table 4.5). Only 36 of these judgements were on segments identified by all coders, that is 9 segments with four annotations each. When we consider leniently matching subsegments, the number of items with four referent segment associations rises to 45, for a total of 180 judgements. Expected disagreement (agreement) is almost maximal (minimal) for these sets when using the binary difference metric thus agreement nears average agreement.

The alternative metric is more flexible than the binary difference distance function, but still suffers from the problem of overestimating chance disagreement by considering coders could choose among more referent segments than are actually possible according to the annotation guidelines. When computing expected agreement or disagreement, the limitations in the choices of segments (past segments by the other speaker) would have to be considered as they decrease agreement and disagreement that could happen by chance. The latter would require the definition of a metric involving the length of the turns and the boundaries of the target segments, making it 0 for segments that belong to the turns of the same speaker or that come

with respect to their combined length taking values between 0 and 1, but results were inconclusive.

later in the dialogue[24].

4.3.3 Aggregating Annotations for the Second Stage

We obtained a single segmented dataset annotated with dialogue act functions and referent segments, aggregating the annotations from the first stage by implementing the method described below:

1. The segment boundaries and their annotations identified by all coders are collected as shown in the example of Figure 4.15(a).

2. Opening and closing segment boundaries are filtered according to an *agreement threshold*, T, between 1 and the total number of coders that can be set when running the algorithm. Only boundaries identified by at least T annotators are kept in the resulting annotation[25].

3. For each one of these boundaries, the annotation labels chosen by the highest number of annotators are kept:

    - If two or more labels meet the criteria above for one of the boundaries in the segment, the ambiguity is resolved by looking at the other boundary.

    - If there is ambiguity in or between the two boundaries of a segment, one label is chosen based on their frequencies. The criterion chosen at runtime can either favour *coverage*, choosing the label

[24] R. Artstein (personal communication, 3 October 2012) expressed concerns about the adequacy of any metric that used the distance between referents as a measure of agreement. He suggested that a distance function capturing semantic content similarities might capture better the similarities and differences between coders' choices.

[25] There were a few cases – 3 to 6 occurrences in the entire corpus, depending on the runtime parameter settings – in which one of the boundaries of a resulting segment is missing. We resolved this by taking the longest stretch of unsegmented speech before or after the unmatched end or start boundary.

with the least frequency, or *conservative* choosing the most frequent one. Frequencies are taken from the distribution of labels in the entire annotation as shown in the Table 4.5.

4. For each remaining segment annotated as a responsive dialogue act, the referent segment chosen by the highest number of coders is chosen.
   - If one or more are candidates, the most recent one is chosen.
   - If the chosen referent segment does not exist in the referred turn (because the boundary did not pass the filter in 2.), the segment which contains the start boundary is chosen instead.

5. The resulting annotation – see the example in Figure 4.15(b) – is written to an XML file and used as input to the second stage.

We generated all 4 variations for the different runtime parameters and decided to take $T = 2$ as the agreement threshold for deciding on what segment boundaries stay and the conservative criterion for when there was not a majority in the annotation of dialogue acts, taking the label with the highest frequency. The value of $T = 3$ for the threshold produced too few segments for the judgments we wanted to obtain in the second stage, so we went with the lower value. A fragment of the resulting dataset is given in Appendix A, Section A.1.2 – the entire dataset is available at `http://mcs.open.ac.uk/nlg/non-cooperation/`. Table 4.15 shows the number of occurrences and the frequency of each dialogue act function.

4.3.4 Second Annotation Stage: Selecting Content Features

In the second stage, annotators received the interview transcripts – segmented and annotated with dialogue act functions and referent segments – that resulted from the aggregation method described above. Section A.1.2 in

Legend. Partial annotations are marked directly on the speech transcript. Segments are boundaries are indicated using square brackets and numbered sequentially. Inside the brackets dialogue act functions are shown according to the following key:

A Annotator 1, **B** Annotator 2, **C** Annotator 3, **D** Annotator 4

1 Init-Inform, **2 Init-InfoReq**, **3 Resp-Inform**, **4 Resp-Accept**, **5 Resp-Reject**

Referent segments are indicated after the dialogue act function using the "**@**" symbol and the number of the segment they point to.

Turn	Spkr.	Annotated Speech
0	IR	**(0)**[**A2B2C2D2** Is Osama Bin Laden your prime suspect? **A2B2C2D2**]
1	IE	**(1)**[**A3@(0)B3@(0)C3@(0)D3@(0)** He is the prime suspect. **A3**] **(2)**[**A1** We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. **A1**] **(3)**[**A3@(0)** And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect. **A3B3R3D3**]
2	IR	**(4)**[**A2B2C2D2** Him alone or anybody else? **A2B2C2D2**]
3	IE	**(5)**[**A1B3@(4)C3@(4)D3@(4)** Well, when we assemble the evidence finally, we will present it to people. **A1B3**] **(6)**[**A3@(4)B3@(4)** But as we have said he is the prime suspect. **A3B3C3D3**]

(a) Collected labels

Turn	Spkr.	Annotated Speech
0	IR	**(0)**[**2** Is Osama Bin Laden your prime suspect?]
1	IE	**(1)**[**3@(0)** He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.]
2	IR	**(2)**[**2** Him alone or anybody else?]
3	IE	**(3)**[**3@(2)** Well, when we assemble the evidence finally, we will present it to people.] **(4)**[**3@(2)** But as we have said he is the prime suspect.]

(b) Result of the aggregation

Figure 4.15: Example of a fragment with annotations aggregated for use as input in the second stage: (a) collecting annotations from all coders into single labels and (b) result of the filtering according to the agreement threshold and coverage criteria.

Appendix A contains the input data corresponding to Interview 1. The data for the entire corpus is available at http://mcs.open.ac.uk/nlg/non-cooperation/.

Table 4.15: Number of occurrences and frequency of dialogue act functions in the annotated corpus resulting from aggregating annotations of the first stage

Interview	Init-Inform	Init-InfoReq	Resp-Inform	Resp-Accept	Resp-Reject	Total
1. Brodie and Blair	1 (0.045)	8 (0.364)	13 (0.591)	0 (0.000)	0 (0.000)	22
2. Green and Miliband	6 (0.333)	6 (0.333)	6 (0.333)	0 (0.000)	0 (0.000)	18
3. O'Reilly and Hartman	7 (0.226)	2 (0.065)	8 (0.258)	4 (0.129)	10 (0.322)	31
4. Paxman and Osborne	1 (0.063)	5 (0.313)	6 (0.375)	2 (0.125)	2 (0.125)	16
5. Pym and Osborne	1 (0.063)	5 (0.313)	10 (0.625)	0 (0.000)	0 (0.000)	16
6. Shaw and Thatcher	4 (0.174)	8 (0.348)	10 (0.435)	0 (0.000)	1 (0.043)	23
Total	20 (0.159)	34 (0.270)	53 (0.421)	6 (0.048)	13 (0.103)	126

Table 4.16: Familiarity of the annotators with the interviews, their contexts and the dialogue participants (second annotation stage)

Interview	Context			Dialogue			Interviewer			Interviewee		
	Y	N	NS	Y	N	NS	Y	N	NS	Y	N	NS
1. Brodie and Blair	7	0	0	3	3	1	4	3	0	7	0	0
2. Green and Miliband	3	4	0	2	4	1	2	5	0	7	0	0
3. O'Reilly and Hartman	3	3	1	3	4	0	3	3	1	3	4	0
4. Paxman and Osborne	4	3	0	3	4	0	7	0	0	6	1	0
5. Pym and Osborne	6	1	0	4	3	0	4	3	0	6	1	0
6. Shaw and Thatcher	5	1	1	1	5	1	3	4	0	7	0	0

Legend. Y: *'yes'*, **N**: *'no'*, **NS**: *'not sure'*.
The figures represent the number of annotators that chose each option.

Annotators

Seven annotators were involved in the annotation of content features: the four coders that took part in the first stage, plus another of the author's supervisors and two native English speakers with no background in linguistics or experience in dialogue analysis. As before, annotators were asked to provide information about their familiarity with the interview, its context and the dialogue participants. A summary of their answers appears in Table 4.16.

Results

Recall from Section 4.2.3 that segments annotated with **Init-Inform**, **Init-InfoReq** and **Resp-Inform** dialogue act functions receive binary judgements on their content referred to as content features. Contrary to the annotation of dialogue act functions, in the second stage, coders identified the content features in the same set of segments. This allowed obtaining inter-annotator agreement values for the three coefficients discussed earlier in this section: Krippendorff's α and the multi-rater versions of Cohen's κ and Scott's π (or Siegel and Castellan's K).

Inter-Annotator Agreement for Content Features. Table 4.17 summarises the number of occurrences and the frequency of each content feature. Of the 126 segments in the corpus, 19 are acceptances or rejections and have no associated content features (cf. Table 4.15). The remaining 107 segments are divided into 20 statements (**Init-Inform**), 34 questions (**Init-InfoReq**) and 53 replies (**Resp-Inform**). The content features in each of these segments were annotated by seven coders, resulting in 140, 238 and 371 judgements, respectively. Table 4.17 shows how these choices are distributed between the two options available for each aspect on which content is judged. For example, of 140 annotations for the 20 **Init-Inform** segments: regarding topicality, 132 found the contents to be on-topic and 8 found them to be off-topic with respect to the subject matter of the interview; regarding objectivity, 56 annotations found the contents objective while 84 found them subjective; regarding accuracy, 127 judged the contents accurate and 13 regarded them as inaccurate; and regarding novelty, 129 judgements found that the contents were new while 11 found that the same contents had been expressed earlier in the dialogue. Similar descriptions applies to the annotations for **Init-InfoReq** and **Resp-Inform** segments. It is worth noting that the distribution of choices for some of the features was particularly skewed towards one of the options (e.g. **On-Topic|Off-Topic** for **Init-Inform** and **Init-InfoReq**, **Accurate|Inaccurate** for **Resp-Inform**). We discuss below the effects of this on the inter-annotator agreement results for such features.

Table 4.17: Number of occurrences and frequency of content features in the annotated corpus

Init-Inform Total: 140		**Init-InfoReq** Total: 238		**Resp-Inform** Total: 371	
On-Topic 132 (0.943)	**Off-Topic** 8 (0.057)	**On-Topic** 235 (0.987)	**Off-Topic** 3 (0.013)	**Relevant** 275 (0.741)	**Irrelevant** 96 (0.259)
Objective 56 (0.4)	**Subjective** 84 (0.6)	**Neutral** 68 (0.286)	**Loaded** 170 (0.714)	**Objective** 208 (0.561)	**Subjective** 163 (0.439)
Accurate 127 (0.907)	**Inaccurate** 13 (0.093)	**Reasonable** 225 (0.945)	**Unreasonable** 13 (0.055)	**Accurate** 365 (0.984)	**Inaccurate** 6 (0.016)
New 129 (0.921)	**Repeated** 11 (0.079)	**New** 211 (0.887)	**Repeated** 27 (0.113)	**New** 288 (0.776)	**Repeated** 83 (0.224)
				Complete 182 (0.491)	**Incomplete** 189 (0.509)

Table 4.18 shows the values of agreement for Krippendorff's α, observed and expected disagreement, observed (or average) agreement A_o, and multi-rater versions of Cohen's κ and Scott's π (or Siegel and Castellan's K) with their respective expected agreements A_e – recall that observed agreement is the same for both coefficients and as given under A_o. We report on agreement for the content features individually, aggregated for each dialogue act function and overall for the entire corpus.

As Table 4.18 shows, agreement varies considerably across features, from "poor" (**Accurate**|**Inaccurate** for **Resp-Inform**) to "perfect" (**New**|**Repeated** for **Init-InfoReq**), using the terminology of Figure 4.12. Overall agreement is "moderate" ($\alpha = 0.454$). Agreement per dialogue act function is also "moderate" for questions ($\alpha = 0.563$) and replies ($\alpha = 0.438$), but falls slightly below the 0.40 threshold and is "fair" for initiating statements ($\alpha = 0.398$).

Looking at the individual features, agreement is consistently high for **New**|**Repeated** for the three dialogue act functions, with values of α between 0.641 ("substantial") and 0.806 ("perfect"). Agreement is "moderate"

Table 4.18: Inter-annotator agreement for content features

Content Feature	α (D_o, D_e)	A_o	κ (A_e)	$\pi\|K$ (A_e)
Init-Inform	0.398 $(0.137, 0.227)$	0.863	0.402 (0.772)	0.393 (0.775)
On-Topic | Off-Topic	0.079 $(0.100, 0.109)$	0.900	0.083 (0.891)	0.072 (0.892)
Objective | Subjective	0.370 $(0.305, 0.483)$	0.695	0.377 (0.510)	0.365 (0.520)
Accurate | Inaccurate	0.467 $(0.090, 0.170)$	0.910	0.467 (0.830)	0.463 (0.832)
New | Repeated	0.641 $(0.052, 0.146)$	0.948	0.640 (0.855)	0.638 (0.855)
Init-InfoReq	0.563 $(0.081, 0.185)$	0.919	0.564 (0.814)	0.560 (0.816)
On-Topic | Off-Topic	0.104 $(0.022, 0.025)$	0.978	0.105 (0.975)	0.100 (0.975)
Neutral | Loaded	0.481 $(0.213, 0.410)$	0.787	0.486 (0.586)	0.478 (0.592)
Reasonable | Unreasonable	0.514 $(0.050, 0.104)$	0.950	0.512 (0.897)	0.512 (0.897)
New | Repeated	0.806 $(0.039, 0.202)$	0.961	0.805 (0.799)	0.805 (0.799)
Resp-Inform	0.438 $(0.198, 0.352)$	0.802	0.443 (0.645)	0.436 (0.649)
Relevant | Irrelevant	0.407 $(0.228, 0.385)$	0.772	0.411 (0.613)	0.405 (0.616)
Objective | Subjective	0.316 $(0.338, 0.494)$	0.662	0.333 (0.493)	0.314 (0.507)
Accurate | Inaccurate	-0.014 $(0.032, 0.032)$	0.968	-0.014 (0.968)	-0.016 (0.968)
New | Repeated	0.763 $(0.083, 0.348)$	0.917	0.762 (0.652)	0.762 (0.653)
Complete | Incomplete	0.383 $(0.309, 0.501)$	0.691	0.385 (0.498)	0.382 (0.500)
Overall	0.454 $(0.143, 0.262)$	0.857	0.458 (0.736)	0.452 (0.739)

for the features **Accurate|Inaccurate** for **Init-Inform**, **Neutral|Loaded** and **Reasonable|Unreasonable** for **Init-InfoReq**, and **Relevant|Irrelevant** for **Resp-Inform**, with values of α between 0.407 and 0.514. Agreement is "fair" for **Objective|Subjective** for **Init-Inform** ($\alpha = 0.370$) and **Resp-Inform** ($\alpha = 0.316$) and for **Complete|Incomplete** for **Resp-Inform** ($\alpha = 0.383$). These results are consistent with the feedback we received from the annotators that relevance, objectivity and completeness were generally hard to assess with the available context and that the assessment depended strongly on specific knowledge which could vary among coders.

The lowest agreement values were for **Accurate|Inaccurate** for **Resp-Inform** ($\alpha = -0.014$), and for **On-Topic|Off-Topic** for **Init-Inform** ($\alpha =$

0.079) and **Init-InfoReq** ($\alpha = 0.104$). Normally, such low agreement would render the categories unreliable. However, these results can be explained by the skewed distributions of judgements towards one of the available options, as pointed out above. In fact, a closer inspection of the agreement values for these three cases shows that average agreement (A_o) was almost perfect: 96.8%, 90% and 97.8%, respectively. The low scores on the agreement coefficients are due to the small number of judgements in one of the options which brought expected disagreement down to, respectively, 0.032, 0.109 and 0.025. This turned the high observed agreement on the most frequent choices insignificant with respect to the slightest disagreement on the rarer alternatives. This means that the data is inconclusive regarding the reliability of judgements on accuracy of replies and topical adequacy of initiating statements and questions.

We checked for outliers by excluding the annotations of each coder and comparing the agreement values above with those of the resulting reduced datasets. Agreement decreased after excluding the annotations of each coder except for annotator 7. Table 4.19 shows the values of agreement for the dataset resulting when excluding the annotations of this coder. As before, agreement values are inconclusive regarding accuracy of replies and topical adequacy of initiating statements and of questions. However, for all other content features, agreement is above the 0.40 threshold and ranges from "moderate" (**Objective**|**Subjective** and **Accurate**|**Inaccurate** for **Init-Inform**, **Neutral**|**Loaded** for **Init-InfoReq**, and **Relevant**|**Irrelevant**, **Objective**|**Subjective** and **Complete**|**Incomplete** for **Resp-Inform**), to "substantial" (**New** | **Repeated** for **Init-Inform** and **Resp-Inform**, **Reasonable**|**Unreasonable** for **Init-InfoReq**), to "perfect" (**New** | **Repeated** for **Init-InfoReq**).

Table 4.19: Inter-annotator agreement for content features (excluding annotations of coder 7)

Content Features	α (D_o, D_e)	A_o	κ (A_e)	$\pi\|K$ (A_e)
Init-Inform	0.440 (0.131, 0.234)	0.869	0.440 (0.767)	0.436 (0.768)
On-Topic | Off-Topic	0.097 (0.113, 0.125)	0.887	0.101 (0.874)	0.089 (0.876)
Objective | Subjective	0.474 (0.263, 0.501)	0.737	0.474 (0.500)	0.470 (0.503)
Accurate | Inaccurate	0.437 (0.087, 0.154)	0.913	0.438 (0.846)	0.433 (0.847)
New | Repeated	0.611 (0.060, 0.154)	0.940	0.610 (0.846)	0.607 (0.847)
Init-InfoReq	0.616 (0.069, 0.180)	0.931	0.616 (0.820)	0.613 (0.822)
On-Topic | Off-Topic	0.125 (0.025, 0.029)	0.975	0.126 (0.971)	0.120 (0.971)
Neutral | Loaded	0.537 (0.175, 0.377)	0.825	0.541 (0.620)	0.535 (0.625)
Reasonable | Unreasonable	0.637 (0.037, 0.103)	0.963	0.635 (0.898)	0.635 (0.898)
New | Repeated	0.812 (0.039, 0.209)	0.961	0.811 (0.792)	0.811 (0.792)
Resp-Inform	0.494 (0.176, 0.349)	0.824	0.495 (0.650)	0.493 (0.652)
Relevant | Irrelevant	0.431 (0.230, 0.404)	0.770	0.434 (0.593)	0.429 (0.597)
Objective | Subjective	0.432 (0.264, 0.465)	0.736	0.435 (0.533)	0.430 (0.537)
Accurate | Inaccurate	−0.010 (0.025, 0.025)	0.975	−0.010 (0.975)	−0.013 (0.975)
New | Repeated	0.765 (0.082, 0.348)	0.918	0.765 (0.653)	0.764 (0.653)
Complete | Incomplete	0.440 (0.281, 0.501)	0.719	0.441 (0.498)	0.439 (0.500)
Overall	0.505 (0.129, 0.261)	0.871	0.506 (0.738)	0.503 (0.740)

4.4 Assessing Cooperation in Annotated Dialogue

This section describes the algorithm for automatically assessing conversational cooperation in annotated dialogue. The method is applied to the annotated corpus of political interviews discussed in the previous section.

4.4.1 Mapping Annotations to Action Labels

As a first step, the dialogue act functions and content features in the annotations are mapped to the **action labels** introduced in Chapter 3 for defining dialogue games. The mapping is carried out automatically, based on rules that are tailored to a specific dialogue game and coding scheme pair.

This approach allows for a separation between the prescriptive nature of the dialogue game and the descriptive character of the coding scheme. Such independence facilitates, for instance, changing the rules of the dialogue game so that it better relates to the social norms, conventions and expectations of different cultural backgrounds, while keeping the coding scheme unchanged and using the same annotated data.

Table 4.20 shows how the dialogue act functions and content features of the coding scheme for political interviews are mapped to the action labels of the political interview dialogue game presented in Chapter 3. As argued in Section 3.5, the validity of contributions with respect to a dialogue game depend on the role of each speaker in the conversation. For this reason, the rules for mapping annotations to action labels are given separately for the interviewer in Table 4.20(a) and for the interviewee in Table 4.20(b).

For the interviewer, an initiating statement is mapped as a valid statement if it is on-topic with respect to the subject matter of the interview and if the information conveyed is objective, accurate and has not been expressed before. This means that segments uttered by the interviewer and annotated as **Init-Inform** dialogue acts with the content features **On-Topic**, **Objective**, **Accurate** and **New** are assigned the `valid-statement` action label. This is specified by the first rule of Table 4.20(a). The second rule states that if *any* of the content features are different from the respective choices, above the **Init-Inform** segment is assigned an `invalid-statement` action label. Similarly, questions are valid if they are on-topic, neutral and reasonable (i.e. if it is in the power of the interviewee to provide an answer)[26]. If any of these conditions are unmet, the question is invalid. These two cases are specified by the third and fourth rules in Table 4.20(a). Any replies from the

[26] Note that it is not a requirement that questions are new as the repetition of a question works as the implicit rejection of a reply. We discuss this further later in the chapter.

Table 4.20: Mapping annotations to action labels in political interviews

Annotation Scheme				Dialogue Game
Dialogue Act		**Content Features**		**Action Label**
Init-Inform	+	**On-Topic** *and* **Objective** *and* **Accurate** *and* **New**	⟶	`valid-statement`
Init-Inform	+	**Off-Topic** *or* **Subjective** *or* **Inaccurate** *or* **Repeated**	⟶	`invalid-statement`
Init-InfoReq	+	**On-Topic** *and* **Neutral** *and* **Reasonable**	⟶	`valid-question`
Init-InfoReq	+	**Off-Topic** *or* **Loaded** *or* **Unreasonable**	⟶	`invalid-question`
Resp-Inform	+	*Any*	⟶	`invalid-reply`
Resp-Accept		⟶		`acceptance`
Resp-Reject		⟶		`rejection`

(a) Interviewer segments

Annotation Scheme				Dialogue Game
Dialogue Act		**Content Features**		**Action Label**
Init-Inform	+	*Any*	⟶	`invalid-statement`[a]
Init-Inform	+	**On-Topic** *and* **Accurate** *and* **New**	⟶	`valid-statement`[b]
Init-Inform	+	**Off-Topic** *or* **Inaccurate** *or* **Repeated**	⟶	`invalid-statement`[b]
Init-InfoReq	+	*Any*	⟶	`invalid-question`
Resp-Inform	+	**Relevant** *and* **Accurate** *and* **New**	⟶	`valid-reply`
Resp-Inform	+	**Irrelevant** *or* **Inaccurate** *or* **Repeated**	⟶	`invalid-reply`
Resp-Accept		⟶		`acceptance`
Resp-Reject		⟶		`rejection`

(b) Interviewee segments

[a] If the interview starts with a question by the interviewer.

[b] In the first turn of an interview that starts with a statement by the interviewee.

interviewer are regarded as invalid[27] so, regardless of the associated content features, segments by the interviewer annotated as **Resp-Inform** are assigned an `invalid-reply` action label. The last two rules in Table 4.20(a) refer to acceptances and rejections. As **Resp-Accept** and **Resp-Reject** dialogue act functions have no associated content features, they are mapped directly to the corresponding action label – respectively, `acceptance` and `rejection`.

For the interviewee, initiating statements convey uninvited information and are generally considered invalid, regardless of their associated content features. This means that, as specified by the first rule in Table 4.20(b), segments spoken by the interviewee and annotated as **Init-Inform** dialogue acts are assigned an `invalid-statement` action label. An exception to this rule is when interviews start with a statement by the interviewee (e.g. Interview 2 in the corpus). In such special cases, the mapping is the same as for initiating statements spoken by the interviewer. This is specified by the second and third rules in Table 4.20(b). Any questions by the interviewee are considered invalid (see Footnote 27 above) which is specified by the fourth rule in the table. Replies are valid if they are relevant with respect to the question they respond to and if the information conveyed is accurate and has not been presented earlier in the conversation. Thus, segments spoken by the interviewee annotated as **Resp-Inform** dialogue acts with content features **Relevant**, **Accurate** and **New** are assigned a `valid-reply` action label. When any of these conditions are unmet, replies are invalid. These mappings are specified by the fifth and sixth rules in Table 4.20(b). It is worth noting that the objectivity of the information provided by the interviewee is not required for the validity of a reply. This is because interviewees are often

[27] Recall from Chapter 3 that according to the dialogue game interviewees are not allowed to ask questions (except for clarification questions which are not part of the current study). Therefore, interviewers should respond to any such questions by rejecting them instead of providing a reply. See turns 13 and 14 in Figure 4.16(a) for an example.

asked to provide their points of view and personal opinions on certain issues and this involve subjective information. **Complete|Incomplete** content features are also not part of the mapping for **Resp-Inform** dialogue acts because a complete answer to a question can take up several segments. This aspect is dealt with at the turn-level when measuring cooperation. As before, acceptances and rejections are mapped directly from annotations to the corresponding action labels.

Figure 4.16 presents an example of the mapping using an excerpt of Interview 6 from the annotated corpus. For each invalid action label, the offending content features are given as reasons for the invalidity. Also, it is shown how all referent segments and the content features **New|Repeated** for **Init-InfoReq** segments and **Complete|Incomplete** for **Resp-Inform** segments are kept for processing in later steps of the method. Below, these will be represented by their initial in bold face after the label name (e.g. `valid-question` **N**), consistently with the structure for action labels presented in Section 3.4.1.

4.4.2 Measuring Cooperation in Dialogue

As stated in Chapter 3, linguistic cooperation of a dialogue participant with respect to a conversational setting equates to the participant following the rules of the dialogue game for that conversational setting. From this perspective, each turn in a dialogue is associated with an amount of cooperation and an amount of non-cooperation, given by the number of dialogue rules, respectively, conformed with and violated in the turn. The instances in which rules are conformed with are called **cooperative features** and those in which rules are broken are called **non-cooperative features**. Recall the following fragment from Chapter 3 (page 79):

Turn	Spkr.	Annotated Speech
6	IR	**(6)[Init-InfoReq ⟨On-Topic, Loaded, Reasonable, New⟩** Following the Falklands War, did hubris from having won that war make you believe that you could persuade the Chinese that Britain should continue administering Hong Kong with an umbrella of Chinese sovereignty?]
7	IE	**(7)[Resp-Inform@(6) ⟨Relevant, Objective, Accurate, New, Complete⟩** No, there was no hubris in Falklands, only a fantastic relief that our people were once again free and we were not going to have an aggressor taking over British land and British people. And we don't like aggression anywhere in the world, that is why we believe in strong defense.]
8	IR	**(8)[Init-InfoReq ⟨On-Topic, Neutral, Reasonable, New⟩** Well, Sir Percy Craddock, Britain's Ambassador to China said that you had to be persuaded, that you had to be told, that there was no way Britain was going to remain an administrative force of Hong Kong with the Chinese being the mere sovereigns.]
9	IE	**(9)[Resp-Inform@(8) ⟨Relevant, Objective, Accurate, New, Complete⟩** Well, that Deng Xiaoping told me. I'll tell you what he told me. I have written it. I said that we have done so well for Hong Kong, for Hong Kong people, that can we not have another lease say for another 50 years? He reacted very quickly. He said no. I said can we not have another lease? I said we have done so well on a territory which I know will eventually return to you. Wouldn't you really let us have, it would be an act of sovereignty to give us a management contract?]
10	IR	**(10)[Init-Inform ⟨On-Topic, Subjective, Accurate, New⟩** They were outraged.]
		(11)[Init-InfoReq ⟨On-Topic, Neutral, Reasonable, New⟩ Is that when Mr. Deng told you that if the Chinese wanted to they could walk right in here and take Hong Kong?]
11	IE	**(12)[Resp-Inform@(11) ⟨Relevant, Objective, Accurate, New, Complete⟩** Oh yes he said he could. But I know that I didn't need to be told. That is why I had to ask him.]
		(13)[Resp-Inform@(11) ⟨Relevant, Subjective, Accurate, New, Complete⟩ But, he said to me, which really rather shook me: I would rather recover Hong Kong poverty stricken than let the British have another period of administration over Hong Kong. Now, that shows you the communist mind, not concerned about the prosperity, about the well being of the people.]
12	IR	**(14)[Init-InfoReq ⟨On-Topic, Neutral, Reasonable, New⟩** You don't trust him, do you?]
13	IE	**(15)[Resp-Inform@(14) ⟨Relevant, Objective, Accurate, New, Complete⟩** I don't trust a communist,]
		(16)[Init-InfoReq ⟨On-Topic, Loaded, Unreasonable, New⟩ do you?]
14	IR	**(17)[Resp-Reject@(16) ⟨⟩** I can't answer that, I am the reporter asking questions.]

(a) Excerpt from an interview segmented and annotated

Turn	Spkr.	Action Labels		
6	IR	**(6)** : `invalid-question`	**New**	{Reason: **Loaded**}
7	IE	**(7)** : `valid-reply@(6)`	**Complete**	
8	IR	**(8)** : `valid-question`	**New**	
9	IE	**(9)** : `valid-reply@(8)`	**Complete**	
10	IR	**(10)** : `invalid-statement`		{Reason: **Subjective**}
		(11) : `valid-question`	**New**	
11	IE	**(12)** : `valid-reply@(11)`	**Complete**	
		(13) : `valid-reply@(11)`	**Complete**	
12	IR	**(14)** : `valid-question`	**New**	
13	IE	**(15)** : `valid-reply@(14)`	**Complete**	
		(16) : `invalid-question`	**New**	{Reason: **Init-InfoReq** by IE}
14	IR	**(17)** : `rejection@(16)`		

(b) The same excerpt with annotations mapped onto action labels

Figure 4.16: Example of the mapping between annotations and action labels

> Participants can break the rules of the game in two ways: (a) by performing a conversational action that is not allowed for their role and (b) by failing to perform an action they were obliged to perform. Instances of (a) are violations of static obligations, which we call **static non-cooperative features**. Instances of (b) are violations of dynamic obligations, which we call **dynamic non-cooperative features**. An analogous distinction is made for cooperative features, called, respectively, **static cooperative features** and **dynamic cooperative features**. The **degree of cooperation** of each dialogue participant is thus the ratio between the number of cooperative features – static and dynamic – and the total number of features of that participant. In general, this value can be obtained for the entire conversation and for any continuous fragments.

In the rest of the section, we revisit the formalisation of these concepts and describe a method to compute the dynamic obligations of participants in each turn of a dialogue. We also explain how to compute static and dynamic cooperative and non-cooperative features and the degree of non-cooperation. Throughout, we illustrate the method using the political interview conversational setting and examples from the annotated corpus.

Computing Dynamic Obligations

Formally, for a dialogue $D = \langle t_1; \ldots ; t_n\rangle$, where each t_i is a turn, we represent dynamic obligations as two sequences $PO_D = \langle PO_{D,0}; PO_{D,1}; \ldots ; PO_{D,n}\rangle$ and $DO_D = \langle DO_{D,1}; \ldots ; DO_{D,n}\rangle$, where each element is a list $\langle o_1, \ldots , o_k\rangle$ of the **obligations pending** after and **discharged** in turn t_i, respectively. $PO_{D,0}$ is also a list with the obligations pending before the dialogue starts.

As in Chapter 3, each obligation is a pair $o_i = (s_i, l_i)$, where s_i is a speaker and l_i is an action label.

Obligations are **updated** in each turn of the dialogue. This means that, for each turn t_i, $PO_{D,i}$ and $DO_{D,i}$ are computed based on $PO_{D,(i-1)}$ by **discharging** existing obligations and **introducing** new pending ones[28]:

- A pending obligation $o = (s, l) \in PO_{D,i-1}$ is **discharged** in turn $t_i = (s_i, L_i)$ if $s = s_i$ and $l_j \succ l$ for some label $l_j \in L_i$. The resulting list of pending obligations $PO_{D,i}$ is $PO_{D,i-1} - \langle o \rangle$. The obligation is added to the list of discharged obligations for the turn so $DO_{D,i}$ becomes $DO_{D,i} \circ \langle o \rangle$. To discharge several obligations in the same turn, we generalise this definition in the obvious way.

- An obligation o is **introduced** in turn $t_i = (s_i, L_i)$ if there is a rule $[(s, l) \rightsquigarrow o]$ in the game such that $s = s_i$ and $l_j \succ l$, for some label $l_j \in L_i$ (meaning that obligations are introduced by implicitly performed actions). After the update, the list of pending obligations is $\langle o \rangle \circ PO_{D,i}$. The list of discharged obligations remains unchanged. To introduce several obligations in the same turn, we generalise this definition in the obvious way.

Those obligations in $PO_{D,(i-1)}$ that are not discharged in turn t_i are carried over to $PO_{D,(i)}$.

Figure 4.18 shows Interview 1 from the annotated corpus with the action labels mapped from the annotations of one of the coders and the dynamic obligations updated after each turn. Let us see how these are computed

[28]In the definition below, recall that the implicit performance relation $\succ$ is reflexive, so for any label l it is $l \succ l$. Also, in the rest of the presentation, we assume that list structures support operations for indexing (denoted as $l[i]$), concatenation ($l_1 \circ l_2$), subtraction ($l_1 - l_2$) and those inherited from sets, such as element membership ($x \in l$) and cardinality ($|l|$).

over a few turns. Figure 4.17 reproduces the dialogue game for political interviews presented at the end of Chapter 3 .

As the interview – here D_1 – starts with the interviewer holding the floor, we assume that at the outset of the dialogue there is a pending obligation on the interviewer to ask a valid question: $PO_{D_1,0} = \langle(\text{IR}, \texttt{valid-question}\ \mathbf{N})\rangle$. In the first turn, $(\text{IR}, \langle(0) : \texttt{valid-question}\ \mathbf{N}\rangle)$, the interviewer asks a new valid question. This action discharges the only obligation in $PO_{D_1,0}$, which is added to $DO_{D_1,1}$, the list of discharged obligations. By Rule (4), it introduces a new obligation for the interviewee to accept the question, so at the end of the turn it is $PO_{D_1,1} = \langle(\text{IE}, \texttt{acceptance@}(0))\rangle$[29] and $DO_{D_1,1} = \langle(\text{IR}, \texttt{valid-question}\ \mathbf{N})\rangle$.

In the second turn, $(\text{IE}, \langle(1) : \texttt{valid-reply@}(0)\ \mathbf{C}\rangle)$, the interviewee provides a valid reply. This implicitly accepts the question due to Rule 18, discharging the only obligation in $PO_{D_1,1}$, which is added to $DO_{D_1,2}$, and introducing the obligation $(\texttt{ie}, \texttt{valid-reply@}(0)\ \mathbf{C})$ by Rule (5), which commits the interviewee to providing a valid and complete reply. This obligation is discharged immediately by the same action and added to the list of discharged obligations. By application of Rule (7), a new pending obligation is introduced for the interviewer to accept the reply. At the end of the turn, it is $PO_{D_1,2} = \langle(\text{IR}, \texttt{acceptance@}(1))\rangle$ and $DO_{D_1,2} = \langle(\text{IE}, \texttt{acceptance@}(0)); (\text{IE}, \texttt{valid-reply@}(0)\ \mathbf{C})\rangle$.

The interviewer asks a new valid question in the third turn, which implicitly accepts the interviewee's reply according to Rule 17 and introduces an obligation on the interviewer to ask a new valid question due to Rule 8, which

[29] Observe that the application of rules involves a mechanism that resembles **unification** in logic programming. Constants such as the label name `valid-question` and the binary flag **N** in Rule 4 are unified with identical constants in the action label that appears in the dialogue. On the other hand, all occurrences of variables such as q in Rule 4 are unified with the value – 0 in this case – in the action label that appears in the dialogue.

Dialogue Game for Political Interviews (Formal)

$$G_{PI} = (Allow_{PI}, Introduce_{PI}, Discharge_{PI})$$

where

$$\begin{aligned}
Allow_{PI} = \{&[\texttt{ir} : \{\texttt{valid-statement}, \texttt{valid-question}, \texttt{acceptance}, \texttt{rejection}\}], && (1)\\
&[\texttt{ie} : \{\texttt{valid-statement}, \texttt{valid-reply}, \texttt{acceptance}, \texttt{rejection}\}]\} && (2)\\
Introduce_{PI} = \{&[(\texttt{ir}, (\texttt{s}) : \texttt{valid-statement}) \rightsquigarrow (\texttt{ie}, \texttt{acceptance@}(\texttt{s}))], && (3)\\
&[(\texttt{ir}, (\texttt{q}) : \texttt{valid-question } \mathbf{N}) \rightsquigarrow (\texttt{ie}, \texttt{acceptance@}(\texttt{q}))], && (4)\\
&[(\texttt{ie}, \texttt{acceptance@}(\texttt{q})) \rightsquigarrow (\texttt{ie}, \texttt{valid-reply@}(\texttt{q})\ \mathbf{C})], && (5)\\
&[(\texttt{ie}, (\texttt{s}) : \texttt{valid-statement}) \rightsquigarrow (\texttt{ir}, \texttt{acceptance@}(\texttt{s}))], && (6)\\
&[(\texttt{ie}, (\texttt{r}) : \texttt{valid-reply@}(\texttt{q})) \rightsquigarrow (\texttt{ir}, \texttt{acceptance@}(\texttt{r}))], && (7)\\
&[(\texttt{ir}, \texttt{acceptance}) \rightsquigarrow (\texttt{ir}, \texttt{valid-question } \mathbf{N})], && (8)\\
&[(\texttt{ir}, (\texttt{s}) : \texttt{invalid-statement}) \rightsquigarrow (\texttt{ie}, \texttt{rejection@}(\texttt{s}))], && (9)\\
&[(\texttt{ir}, (\texttt{q}) : \texttt{invalid-question}) \rightsquigarrow (\texttt{ie}, \texttt{rejection@}(\texttt{q}))], && (10)\\
&[(\texttt{ir}, (\texttt{r}) : \texttt{invalid-reply}) \rightsquigarrow (\texttt{ie}, \texttt{rejection@}(\texttt{r}))], && (11)\\
&[(\texttt{ie}, (\texttt{s}) : \texttt{invalid-statement}) \rightsquigarrow (\texttt{ir}, \texttt{rejection@}(\texttt{s}))], && (12)\\
&[(\texttt{ie}, (\texttt{q}) : \texttt{invalid-question}) \rightsquigarrow (\texttt{ir}, \texttt{rejection@}(\texttt{q}))], && (13)\\
&[(\texttt{ie}, (\texttt{r}) : \texttt{invalid-reply}) \rightsquigarrow (\texttt{ir}, \texttt{rejection@}(\texttt{r}))]\} && (14)\\
Discharge_{PI} = \{&[*\texttt{-question } \mathbf{R} \succ \texttt{rejection}], && (15)\\
&[*\texttt{-statement} \succ \texttt{acceptance}], && (16)\\
&[*\texttt{-question } \mathbf{N} \succ \texttt{acceptance}], && (17)\\
&[*\texttt{-reply} \succ \texttt{acceptance}]\} && (18)
\end{aligned}$$

Figure 4.17: Dialogue game for political interviews (repeated)

it also discharges, plus one obligation on the interviewee to accept the question, according to Rule 4. The list of pending obligations at the end of the turn is then $PO_{D_1,3} = \langle(\text{IE}, \texttt{acceptance@}(2))\rangle$. The obligations discharged in the turn are $DO_{D_1,3} = \langle(\text{IR}, \texttt{acceptance@}(1)); (\text{IR}, \texttt{valid-question } \mathbf{N})\rangle$

Turn 4 shows an instance in which the question is implicitly accepted, which introduces an obligation on the interviewee to provide a complete valid reply. This is done by providing a valid incomplete reply, which is followed by an invalid reply, $(\text{IE}, \langle(3) : \texttt{valid-reply@}(2)\ \mathbf{I}; (4) : \texttt{invalid-reply@}(2)\ \mathbf{I}\rangle)$. These actions introduce obligations on the interviewer to accept the valid – yet incomplete – reply and to reject the invalid reply. They

Turn	Spkr.	Action Labels	Pending Obligations	Discharged Obligations
0			IR: valid-question **N**	
1	IR	**(0)**:valid-question **N**	IE: acceptance@(0)	IR: valid-question **N**
2	IE	**(1)**:valid-reply@(0) **C**	IR: acceptance@(1)	IE: acceptance@(0), IE: valid-reply@(0) **C**
3	IR	**(2)**:valid-question **N**	IE: acceptance@(2)	IR: acceptance@(1), IR: valid-question **N**
4	IE	**(3)**:valid-reply@(2) **I** **(4)**:invalid-reply@(2) **I**	IR: acceptance@(3), IR: rejection@(4), IE: valid-reply@(2) **C**	IE: acceptance@(2)
5	IR	**(5)**:valid-question **N**	IE: acceptance@(5), IR: rejection@(4), IE: valid-reply@(2) **C**	IR: acceptance@(3), IR: valid-question **N**
6	IE	**(6)**:valid-reply@(5) **C** **(7)**:invalid-reply@(5) **C** **(8)**:invalid-reply@(5) **C**	IR: acceptance@(6), IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IE: acceptance@(5), IE: valid-reply@(5) **C**
7	IR	**(9)**:valid-question **N**	IE: acceptance@(9), IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IR: acceptance@(6), IR: valid-question **N**
8	IE	**(10)**:invalid-reply@(9) **I**	IR: rejection@(10), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IE: acceptance@(9)
9	IR	**(11)**:valid-question **R**	IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IR: rejection@(10)
10	IE	**(12)**:invalid-reply@(11) **I**	IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	—
11	IR	**(13)**:valid-question **N**	IE: acceptance@(13), IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IR: valid-question **N**
12	IE	**(14)**:valid-reply@(13) **C** **(15)**:invalid-reply@(13) **C**	IR: acceptance@(14), IR: rejection@(15), IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IE: acceptance@(13), IE: valid-reply@(13) **C**
13	IR	**(16)**:valid-statement **(17)**:valid-question **N**	IE: acceptance@(16), IE: acceptance@(17), IR: rejection@(15), IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IR: acceptance@(14), IR: valid-question **N**
14	IE	**(18)**:valid-reply@(17) **I** **(19)**:valid-reply@(17) **C**	IR: acceptance@(18), IR: acceptance@(19), IR: rejection@(15), IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IE: acceptance@(16), IE: acceptance@(17), IE: valid-reply@(17) **C**
15	IR	**(20)**:valid-question **N**	IE: acceptance@(20), IR: rejection@(15), IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IR: acceptance@(18), IR: acceptance@(19), IR: valid-question **N**
16	IE	**(21)**:valid-reply@(20) **I**	IE: valid-reply@(20) **C**, IR: rejection@(15), IR: rejection@(12), IE: valid-reply@(9) **C**, IR: rejection@(7), IR: rejection@(8), IR: rejection@(4), IE: valid-reply@(2) **C**	IE: acceptance@(20)

Content features: **(N)ew**\|**(R)epeated**, **(C)omplete**\|**(I)ncomplete**.

Figure 4.18: Action labels and dynamic obligations for Interview 1

do not discharge the obligation on the interviewer to provide a complete valid reply, so the obligations pending at the end of the turn are $PO_{D_1,4} = \langle(\textsc{ir}, \texttt{acceptance}@(3)); (\textsc{ir}, \texttt{rejection}@(3)); (\textsc{ie}, \texttt{valid-reply}@(2))\ \mathbf{C}\rangle$ and the obligations discharged in the turn are $DO_{D_1,4} = \langle(\textsc{ie}, \texttt{acceptance}@(2))\rangle$.

Moving forward to Turns 7 through 9, it can be observed how the repetition of a question works as a rejection. The new valid question in Turn 7 introduces the obligation $(\textsc{ie}, \texttt{acceptance}@(9))$, for the interviewee to accept the question. The acceptance is implicitly performed in Turn 8 by means of an invalid reply which introduces an obligation on the interviewee to provide a complete valid reply, $(\textsc{ie}, \texttt{valid-reply}@(9)\ \mathbf{C})$. The invalid reply introduces an obligation on the interviewer to reject it, $(\textsc{ir}, \texttt{rejection}@(10))$. The repeated question in Turn 9 implicitly performs this rejection discharging the obligation and inviting a complete valid reply from the interviewee.

Turn 12 illustrates the treatment of complete valid replies that extend over several actions. The interviewee provides an incomplete valid reply, followed by a valid reply that completes the answer. This discharges the obligation introduced implicitly by the first action and introduces a new obligation of acceptance on the interviewer.

Algorithm 1 describes the procedure for automatically computing the sequences $PO_D = \langle PO_{D,1}; \ldots; PO_{D,n}\rangle$ and $DO_D = \langle DO_{D,1}; \ldots; DO_{D,n}\rangle$ of pending and discharged dynamic obligations for a dialogue D and starting obligations $PO_{D,0}$, given dialogue game G. The implementation of this algorithm is actually more complicated as it requires that responsive implicitly performed actions, such as acceptances and rejections, be matched to the actions they refer to. The criterion we followed considers any actions that need be accepted or rejected by the current speaker and for which the corresponding obligations were introduced in the turn immediately preceding the current one. The implicit action performed with respect to such actions is then determined by the rules in the set *Discharge* of the dialogue game[30].

Algorithm 1 Computing dynamic obligations for dialogue D, initial pending obligations $PO_D[0]$ and dialogue game G.

```
for i in {1... length(D)} do                                    [for each turn in the dialogue...]

  (speaker, labels) ← D[i]                                       [take actions in current turn]
  pending-obligations ← PO_D[i − 1]                          [take previous pending obligations]

                                  [introduce new obligations]
  for label in labels do                                         [for each action in the turn...]
    for rule in G(2) do                                [for each rule in the set Introduce...]
      [(rule-speaker, rule-label) ⇝ obligation] ← rule
      if (speaker = rule-speaker) ∧ (label ≻ rule-label) then
        pending-obligations ← ⟨obligation⟩ ∘ pending-obligations
      end if
    end for
  end for

                               [discharge obligations met in turn]
  discharged-obligations ← ⟨⟩
  for label in labels do                                         [for each action in the turn...]
    for obligation in pending-obligations do
      (obligation-speaker, obligation-label) ← obligation
      if (speaker = obligation-speaker) ∧ (label ≻ obligation-label) then
        pending-obligations ← pending-obligations − ⟨obligation⟩
        discharged-obligations ← discharged-obligations ∘ ⟨obligation⟩
      end if
    end for
  end for

  PO_D[i] ← pending-obligations                              [set obligations pending after turn]
  DO_D[i] ← discharged-obligations                         [set obligations discharged in turn]
end for
```

Computing Cooperative and Non-Cooperative Features

As introduced in the previous chapter, cooperative and non-cooperative features are instances of, respectively, observed and neglected static and dy-

[30]Explicitly discharged obligations are straightforward to deal with by following the unification-like binding of variables and constants discussed earlier and bearing in mind that the performance relation, $\succ$, is reflexive.

namic obligations:

> **Static cooperative features:** actions performed by the speaker that are allowed for by his or her role in the dialogue.
>
> **Static non-cooperative features:** actions performed by the speaker that are disallowed for by his or her role in the dialogue.
>
> **Dynamic cooperative features:** obligations on the speaker that were discharged in the current turn.
>
> **Dynamic non-cooperative features:** obligations on the speaker that were not discharged in the current turn.

Formally, for a dialogue $D = \langle t_1; \ldots; t_n \rangle$ features will be grouped in two sequences, $SF_D = \langle sf_1; \ldots; sf_n \rangle$ and $DF_D = \langle df_1; \ldots; df_n \rangle$, of static and dynamic features, respectively. The elements in both sequences are triples (s_i, C_i, NC_i) where s_i is the speaker in turn t_i, C_i is the list of cooperative features and NC_i is the list of non-cooperative features (static for sf_i and dynamic for df_i). In the rest of the section we show how compute these features. Figure 4.19 reproduces the example in Figure 4.18 with two new columns showing the static and dynamic features. For each turn, cooperative features appear decorated with a '✓' sign and non-cooperative features appear with a '$\times$' sign.

Computing Static Features. In each turn, we check whether the actions performed by the speaker are allowed for his or her role as specified in the dialogue game. If an action is in the the speaker's set of allowed actions, then it constitutes a static cooperative feature, otherwise it becomes a static non-cooperative feature. Formally, this means that in turn $t_i = (s_i, L_i)$, for

Turn	Spkr.	Action Labels	Pending Obligations	Static Features	Dynamic Features
0			IR: valid-question **N**		
1	IR	**(0)**:valid-question **N**	IE: acceptance@(0)	✓valid-question **N**	✓valid-question **N**
2	IE	**(1)**:valid-reply@(0) **C**	IR: acceptance@(1)	✓valid-reply@(0) **C**	✓acceptance@(0) ✓valid-reply@(0) **C**
3	IR	**(2)**:valid-question **N**	IE: acceptance@(2)	✓valid-question **N**	✓acceptance@(1) ✓valid-question **N**
4	IE	**(3)**:valid-reply@(2) **I** **(4)**:invalid-reply@(2) **I**	IR: acceptance@(3) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-reply@(2) **I** ×invalid-reply@(2) **I**	✓acceptance@(2) ×valid-reply@(2) **C**
5	IR	**(5)**:valid-question **N**	IE: acceptance@(5) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-question **N**	✓acceptance@(3) ✓valid-question **N** ×rejection@(4)
6	IE	**(6)**:valid-reply@(5) **C** **(7)**:invalid-reply@(5) **C** **(8)**:invalid-reply@(5) **C**	IR: acceptance@(6) IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-reply@(5) **C** ×invalid-reply@(5) **C** ×invalid-reply@(5) **C**	✓acceptance@(5) ✓valid-reply@(5) **C** ×valid-reply@(2) **C**
7	IR	**(9)**:valid-question **N**	IE: acceptance@(9) IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-question **N**	✓acceptance@(6) ✓valid-question **N** ×rejection@(7) ×rejection@(8) ×rejection@(4)
8	IE	**(10)**:invalid-reply@(9) **I**	IR: rejection@(10) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	×invalid-reply@(9) **I**	✓acceptance@(9) ×valid-reply@(9) **C** ×valid-reply@(2) **C**
9	IR	**(11)**:valid-question **R**	IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-question **R**	✓rejection@(10) ×rejection@(7) ×rejection@(8) ×rejection@(4)
10	IE	**(12)**:invalid-reply@(11) **I**	IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	×invalid-reply@(11) **I**	×valid-reply@(9) **C** ×valid-reply@(2) **C**
11	IR	**(13)**:valid-question **N**	IE: acceptance@(13) IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-question **N**	✓valid-question **N** ×rejection@(12) ×rejection@(7) ×rejection@(8) ×rejection@(4)
12	IE	**(14)**:valid-reply@(13) **C** **(15)**:invalid-reply@(13) **C**	IR: acceptance@(14) IR: rejection@(15) IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-reply@(13) **C** ×invalid-reply@(13) **C**	✓acceptance@(13) ✓valid-reply@(13) **C** ×valid-reply@(9) **C** ×valid-reply@(2) **C**
13	IR	**(16)**:valid-statement **(17)**:valid-question **N**	IE: acceptance@(16) IE: acceptance@(17) IR: rejection@(15) IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-statement ✓valid-question **N**	✓acceptance@(14) ✓valid-question **N** ×rejection@(15) ×rejection@(12) ×rejection@(7) ×rejection@(8) ×rejection@(4)
14	IE	**(18)**:valid-reply@(17) **I** **(19)**:valid-reply@(17) **C**	IR: acceptance@(18) IR: acceptance@(19) IR: rejection@(15) IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-reply@(17) **I** ✓valid-reply@(17) **C**	✓acceptance@(16) ✓acceptance@(17) ✓valid-reply@(17) **C** ×valid-reply@(9) **C** ×valid-reply@(2) **C**
15	IR	**(20)**:valid-question **N**	IE: acceptance@(20) IR: rejection@(15) IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-question **N**	✓acceptance@(18) ✓acceptance@(19) ✓valid-question **N** ×rejection@(15) ×rejection@(12) ×rejection@(7) ×rejection@(8) ×rejection@(4)
16	IE	**(21)**:valid-reply@(20) **I**	IE: valid-reply@(20) **C**, IR: rejection@(15) IR: rejection@(12) IE: valid-reply@(9) **C** IR: rejection@(7) IR: rejection@(8) IR: rejection@(4) IE: valid-reply@(2) **C**	✓valid-reply@(20) **I**	✓acceptance@(20) ×valid-reply@(20) **C** ×valid-reply@(9) **C** ×valid-reply@(2) **C**

Figure 4.19: Static and dynamic cooperative (✓) and non-cooperative (×) features for Interview 1

each $l \in L_i$ we check whether $l \in L$, with $[s_i : L]$ a dialogue game rule in $G(1) = Allow$. If this is the case, then l is added to C_i in $sf_i = (s_i, C_i, NC_i)$. Otherwise, l is added to NC_i.

In the example of Figure 4.19, $t_1 = (\textsc{ir}, \langle(0) : \texttt{valid-question}\ \mathbf{N}\rangle)$ is the first turn. Rule (1) of G_{PI} (see Figure 4.17) is $[\texttt{ir} : \{\texttt{valid-statement}, \texttt{valid-question}, \texttt{acceptance}, \texttt{rejection}\}]$, so this action is a static cooperative feature, and $sf_1 = (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)$ are the static features for this turn. The second turn is $t_2 = (\textsc{ie}, \langle(1) : \texttt{valid-reply@}(0)\ \mathbf{C}\rangle)$. Rule (2) of the dialogue game, $[\texttt{ie} : \{\texttt{valid-statement}, \texttt{valid-reply}, \texttt{acceptance}, \texttt{rejection}\}]$, so $sf_2 = (\texttt{ie}, \langle\texttt{valid-reply@}(0)\ \mathbf{C}\rangle, \langle\rangle)$ are the static features for this turn. The third turn is analogous to Turn 1 and so are its static features. The fourth turn, however, is $t_4 = (\textsc{ie}, \langle(3) : \texttt{valid-reply@}(2)\ \mathbf{I}; (4) : \texttt{invalid-reply@}(2)\ \mathbf{I}\rangle)$. The first action is among those specified for the interviewee's role, but the second one is not and constitutes a static non-cooperative feature. The static features for the fourth turn are thus $sf_4 = (\texttt{ie}, \langle\texttt{valid-reply@}(2)\ \mathbf{I}\rangle, \langle\texttt{invalid-reply@}(2)\ \mathbf{I}\rangle)$. Following this method for the rest of the turns produces the sequence of static features for the entire dialogue:

$$
\begin{aligned}
SF_{D_1} = \langle\, & (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; & (1)\\
& (\texttt{ie}, \langle\texttt{valid-reply@}(0)\ \mathbf{C}\rangle, \langle\rangle)\,; & (2)\\
& (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; & (3)\\
& (\texttt{ie}, \langle\texttt{valid-reply@}(2)\ \mathbf{I}\rangle, \langle\texttt{invalid-reply@}(2)\ \mathbf{I}\rangle)\,; & (4)\\
& (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; & (5)\\
& (\texttt{ie}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\texttt{invalid-reply@}(5)\ \mathbf{C}; \texttt{invalid-reply@}(5)\ \mathbf{C}\rangle)\,; & (6)\\
& (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; & (7)\\
& (\texttt{ie}, \langle\rangle, \langle\texttt{invalid-reply@}(9)\ \mathbf{I}\rangle)\,; & (8)\\
& (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{R}\rangle, \langle\rangle)\,; & (9)\\
& (\texttt{ie}, \langle\rangle, \langle\texttt{invalid-reply@}(11)\ \mathbf{I}\rangle)\,; & (10)\\
& (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; & (11)\\
& (\texttt{ie}, \langle\texttt{valid-reply@}(13)\ \mathbf{C}\rangle, \langle\texttt{invalid-reply@}(13)\ \mathbf{C}\rangle)\,; & (12)\\
& (\texttt{ir}, \langle\texttt{valid-statement}; \texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; & (13)
\end{aligned}
$$

$$(\texttt{ie}, \langle\texttt{valid-reply@}(17)\ \mathbf{I}; \texttt{valid-reply@}(17)\ \mathbf{C}\rangle, \langle\rangle)\,; \tag{14}$$
$$(\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)\,; \tag{15}$$
$$(\texttt{ie}, \langle\texttt{valid-reply@}(20)\ \mathbf{I}\rangle, \langle\rangle)\rangle \tag{16}$$

Algorithm 2 shows the procedure for computing static features in a dialogue D given a dialogue game G.

Computing Dynamic Features. In each turn, we look at the speaker's obligations pending after and discharged in that turn. If an obligation on the speaker has been discharged within the turn, then it constitutes a dynamic cooperative feature, otherwise it becomes a dynamic non-cooperative feature. Formally, this means that for turn $t_i = (s_i, L_i)$, an obligation

Algorithm 2 Computing static features for dialogue D and game G.

for i **in** $\{1 \ldots \mathsf{length}(D)\}$ **do** `[for each turn in the dialogue...]`

  $(speaker, labels) \leftarrow D[i]$ `[take current turn]`

  `[collect actions allowed for the speaker]`

  $allowed\text{-}actions \leftarrow \{\}$
  for $rule$ **in** $G(1)$ **do** `[for each rule in the set Allow...]`
    $[rule\text{-}speaker : rule\text{-}labels] \leftarrow rule$
    if $speaker = rule\text{-}speaker$ **then**
      $allowed\text{-}actions \leftarrow allowed\text{-}actions \cup rule\text{-}labels$
    end if
  end for

  `[compute static features for current turn]`

  $cooperative \leftarrow \langle\rangle$
  $non\text{-}cooperative \leftarrow \langle\rangle$
  for $label$ **in** $labels$ **do** `[for each action in the turn...]`
    if $label.name \in allowed\text{-}actions$ **then**
      $cooperative \leftarrow cooperative \circ \langle label\rangle$
    else
      $non\text{-}cooperative \leftarrow non\text{-}cooperative \circ \langle label\rangle$
    end if
  end for

  `[set static features for current turn]`

  $SF_D[i] \leftarrow (speaker, cooperative, non\text{-}cooperative)$
end for

$o = (s_o, l_o) \in DO_{D,i}$ discharged in the current turn for which $s_o = s_i$ is the current speaker is a dynamic cooperative feature. The action label l_o is thus added to the list C_i of dynamic cooperative features in $df_i = (s_i, C_i, NC_i)$. On the other hand, a pending obligation $o = (s_o, l_o) \in PO_{D,i}$, not discharged in the current turn for which $s_o = s_i$ is the turn speaker is a dynamic non-cooperative feature. The action label l_o is thus added to NC_i of dynamic non-cooperative features in $df_i = (s_i, C_i, NC_i)$.

After the first turn of the example, the list of pending obligations is $PO_{D_1,1} = \langle(\textsc{ie}, \texttt{acceptance@}(0))\rangle$. The list of obligations discharged in the turn is $DO_{D_1,1} = \langle(\textsc{ir}, \texttt{valid-question}\ \mathbf{N})\rangle$. As the speaker is the interviewer, $s_1 = \texttt{ir}$, the dynamic features are $df_1 = (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle)$ for Turn 1.

Turns 2 and 3 also have only cooperative dynamic features. On the other hand, after the forth turn, the list of pending obligations is $PO_{D_1,4} = \langle(\textsc{ir}, \texttt{acceptance@}(3)); (\textsc{ir}, \texttt{rejection@}(3)); (\textsc{ie}, \texttt{valid-reply@}(2))\ \mathbf{C}\rangle$ and the obligations discharged in the turn are $DO_{D_1,4} = \langle(\textsc{ie}, \texttt{acceptance@}(2))\rangle$. As the speaker is the interviewee, $s_4 = \texttt{ie}$, the dynamic features for Turn 4 are $df_4 = (\texttt{ie}, \langle\rangle, \langle\texttt{valid-reply@}(2)\ \mathbf{C}\rangle)$.

Following this method for the rest of the turns produces the sequence of dynamic features for dialogue D_1:

$$
\begin{aligned}
DF_{D_1} = \langle & (\texttt{ir}, \langle\texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle); && (1)\\
& (\texttt{ie}, \langle\texttt{acceptance@}(0); \texttt{valid-reply@}(0)\ \mathbf{C}\rangle, \langle\rangle); && (2)\\
& (\texttt{ir}, \langle\texttt{acceptance@}(1); \texttt{valid-question}\ \mathbf{N}\rangle, \langle\rangle); && (3)\\
& (\texttt{ie}, \langle\texttt{acceptance@}(2)\rangle, \langle\texttt{valid-reply@}(2)\ \mathbf{C}\rangle); && (4)\\
& (\texttt{ir}, \langle\texttt{acceptance@}(3); \texttt{valid-question}\ \mathbf{N}\rangle, \langle\texttt{rejection@}(4)\rangle); && (5)\\
& (\texttt{ie}, \langle\texttt{acceptance@}(5); \texttt{valid-reply@}(5)\ \mathbf{C}\rangle, \langle\texttt{valid-reply@}(2)\ \mathbf{C}\rangle); && (6)\\
& (\texttt{ir}, \langle\texttt{acceptance@}(6); \texttt{valid-question}\ \mathbf{N}\rangle,\\
& \quad \langle\texttt{rejection@}(7); \texttt{rejection@}(8); \texttt{rejection@}(4)\rangle); && (7)\\
& (\texttt{ie}, \langle\texttt{acceptance@}(9)\rangle, \langle\texttt{valid-reply@}(9)\ \mathbf{C}; \texttt{valid-reply@}(2)\ \mathbf{C}\rangle); && (8)\\
& (\texttt{ir}, \langle\texttt{rejection@}(10)\rangle, \langle\texttt{rejection@}(7); \texttt{rejection@}(8); \texttt{rejection@}(4)\rangle); && (9)\\
& (\texttt{ie}, \langle\rangle, \langle\texttt{valid-reply@}(9)\ \mathbf{C}; \texttt{valid-reply@}(2)\ \mathbf{C}\rangle); && (10)
\end{aligned}
$$

$$(\texttt{ir}, \langle \texttt{valid-question}\ \mathbf{N} \rangle,$$
$$\langle \texttt{rejection@}(12); \texttt{rejection@}(7); \texttt{rejection@}(8); \texttt{rejection@}(4) \rangle); \qquad (11)$$
$$(\texttt{ie}, \langle \texttt{acceptance@}(13); \texttt{valid-reply@}(13)\ \mathbf{C} \rangle,$$
$$\langle \texttt{valid-reply@}(9)\ \mathbf{C}; \texttt{valid-reply@}(2)\ \mathbf{C} \rangle); \qquad (12)$$
$$(\texttt{ir}, \langle \texttt{acceptance@}(14); \texttt{valid-question}\ \mathbf{N} \rangle,$$
$$\langle \texttt{rejection@}(15); \texttt{rejection@}(12); \texttt{rejection@}(7);$$
$$\texttt{rejection@}(8); \texttt{rejection@}(4) \rangle); \qquad (13)$$
$$(\texttt{ie}, \langle \texttt{acceptance@}(16); \texttt{acceptance@}(17); \texttt{valid-reply@}(17)\ \mathbf{C} \rangle,$$
$$\langle \texttt{valid-reply@}(9)\ \mathbf{C}; \texttt{valid-reply@}(2)\ \mathbf{C} \rangle); \qquad (14)$$
$$(\texttt{ir}, \langle \texttt{acceptance@}(18); \texttt{acceptance@}(19); \texttt{valid-question}\ \mathbf{N} \rangle,$$
$$\langle \texttt{rejection@}(15); \texttt{rejection@}(12); \texttt{rejection@}(7);$$
$$\texttt{rejection@}(8); \texttt{rejection@}(4) \rangle); \qquad (15)$$
$$(\texttt{ie}, \langle \texttt{valid-reply@}(20)\ \mathbf{C} \rangle,$$
$$\langle \texttt{acceptance@}(20); \texttt{valid-reply@}(9)\ \mathbf{C}; \texttt{valid-reply@}(2)\ \mathbf{C} \rangle)\rangle \qquad (16)$$

Algorithm 3 shows a procedure for computing dynamic features in dialogue D, given pending dynamic obligations PO_D and discharged dynamic obligations DO_D.

Computing the Degree of Non-Cooperation

Once we have computed the static and dynamic features for each turn, we can regard the proportion of these that are cooperative as an indicator of the extent to which each participant acted within the rules of the game. This is the **degree of cooperation** of a dialogue participant with respect to a dialogue game. Formally, for speaker s and dialogue $D = \langle t_1; \ldots; t_n \rangle$ this is:

$$dc_{D,s} = \frac{cf_{D,s}}{cf_{D,s} + ncf_{D,s}}$$

where $cf_{D,s}$ is the number of cooperative features – both static and dynamic – of participant s and $ncf_{D,s}$ is the analogous for non-cooperative features.

This is[31]:

$$cf_{D,s} = \sum_{\substack{i=1 \\ [s_i=s]}}^{n} |sf_i(2)| + |df_i(2)|$$

$$ncf_{D,s} = \sum_{\substack{i=1 \\ [s_i=s]}}^{n} |sf_i(3)| + |df_i(3)|$$

Note that, although these definitions are here expressed for the complete dialogue, the same applies to any contiguous subsequences of turns.

The **degree of non-cooperation** of a dialogue participant s in dialogue D is:

$$dnc_{D,s} = 1 - dc_{D,s} = \frac{ncf_{D,s}}{cf_{D,s} + ncf_{D,s}}$$

Table 4.21 summarises the values involved in computing the degree of non-cooperation for both participants in Interview 1. A printout of the output produced by the Java program implementing the method described above run on the annotation data for Interview 1 is given in Appendix B.

Before moving on to describing how we evaluated the validity of the method, it is worth pointing out one possible shortcoming. As presented above, dynamic non-cooperative features, that is unmet obligations on the current speaker, are never forgotten. This means that they are counted towards the degree of non-cooperation in every turn of the speaker from the turn in which they are first introduced until they are discharged, or until the dialogue ends. Although this is technically reasonable as the rules of the game are actually violated every time a speaker fails to meet an obligation, human observers are more likely to sanction early misbehaviours increasingly less harsh later on in the dialogue. For the validity analysis, in which we

[31] Recall that the elements in the sequences of both static and dynamic features $SF_D = \langle sf_1; \ldots; sf_n \rangle$ and $DF_D = \langle df_1; \ldots; df_n \rangle$ are triples (s_i, C_i, NC_i), where s_i is the speaker in turn t_i, and C_i and NC_i are the associated sequences of, respectively, cooperative and non-cooperative features.

Algorithm 3 Computing dynamic features for dialogue D, given pending dynamic obligations PO_D and discharged dynamic obligations DO_D.

for i **in** $\{1 \ldots \mathsf{length}(D)\}$ **do** `[for each turn in the dialogue...]`

  $(speaker, label) \leftarrow D[i]$ `[take speaker of current turn]`

  `[compute dynamic features for current turn]`

  $cooperative \leftarrow \langle\rangle$
  $non\text{-}cooperative \leftarrow \langle\rangle$

  `[collect obligations on speaker met in current turn]`

  $met\text{-}obligations \leftarrow DO_D[i]$
  for $obligation$ **in** $met\text{-}obligations$ **do**
    $(obligation\text{-}speaker, obligation\text{-}label) \leftarrow obligation$
    $cooperative \leftarrow cooperative \circ \langle obligation\text{-}label) \rangle$
  end for

  `[collect obligations on speaker pending after current turn]`

  $unmet\text{-}obligations \leftarrow PO_D[i]$
  for $obligation$ **in** $unmet\text{-}obligations$ **do**
    $(obligation\text{-}speaker, obligation\text{-}label) \leftarrow obligation$
    if $obligation\text{-}speaker = speaker$ **then**
      $non\text{-}cooperative \leftarrow non\text{-}cooperative \circ \langle obligation\text{-}label \rangle$
    end if
  end for

  `[set dynamic features for current turn]`

  $DF_D[i] \leftarrow (speaker, cooperative, non\text{-}cooperative)$
end for

Table 4.21: Degree of non-cooperation for the participants in Interview 1

	Interviewer	**Interviewee**
Static Features		
Cooperative	9	7
Non-Cooperative	0	6
Dynamic Features		
Cooperative	14	12
Non-Cooperative	21	13
$cf_{D,s}$	23	19
$ncf_{D,s}$	21	19
$\mathbf{dc_{D,s}}$	**0.523**	**0.5**
$\mathbf{dnc_{D,s}}$	**0.477**	**0.5**

contrast the results from the method to human assessment of the speakers' behaviour, we introduced a runtime parameter that causes the method to forget dynamic non-cooperative features after a certain number of turns have passed since they were first introduced. We called this parameter the **dialogue history threshold**.

4.5 Evaluation of the Method (Part 2): a survey study

This section describes a survey study designed for evaluating the **validity** of the method described above. The goal of the study is to analyse how the degree of cooperation resulting from the application of the method to a corpus of interviews correlates to human judgement on the behaviour of the dialogue participants in transcripts of the same dataset (see Figure 4.3). Human observers were asked to base their judgements on the intuitions they have regarding how participants ought to behave in a political interview. Judgements on each dialogue participant were then aggregated and the result checked for correlation with the respective degree of cooperation resulting from the annotated corpus by application of the algorithms described in the previous section.

4.5.1 Degree of Cooperation of the Dialogues in the Corpus

We start by reporting the results of applying the method for automatically computing the degree of cooperation on the annotated corpus. As noted at the end of Section 4.3.4, we only considered the annotations of six of the seven annotators as this set showed increased reliability results. Table 4.22 summarises the results, shown as a bar chart in Figure 4.20.

Table 4.22: Degree of cooperation for the interviews in the annotated corpus

Interview	Mean	Annot.1	Annot.2	Annot.3	Annot.4	Annot.5	Annot.6
1. Brodie and Blair							
Interviewer	0.529	0.412	0.523	0.500	0.571	0.468	0.703
Interviewee	0.515	0.525	0.500	0.500	0.526	0.415	0.625
2. Green and Miliband							
Interviewer	0.498	0.440	0.480	0.440	0.520	0.560	0.545
Interviewee	0.310	0.270	0.297	0.270	0.314	0.412	0.294
3. O'Reilly and Hartman							
Interviewer	0.309	0.313	0.292	0.292	0.292	0.333	0.333
Interviewee	0.671	0.700	0.667	0.667	0.600	0.759	0.633
4. Paxman and Osborne							
Interviewer	0.233	0.407	0.195	0.195	0.167	0.167	0.265
Interviewee	0.239	0.276	0.207	0.276	0.138	0.138	0.4
5. Pym and Osborne							
Interviewer	0.336	0.538	0.448	0.093	0.167	0.419	0.353
Interviewee	0.304	0.400	0.345	0.114	0.182	0.385	0.400
6. Shaw and Thatcher							
Interviewer	0.407	0.478	0.525	0.245	0.300	0.465	0.429
Interviewee	0.320	0.377	0.339	0.271	0.254	0.306	0.370

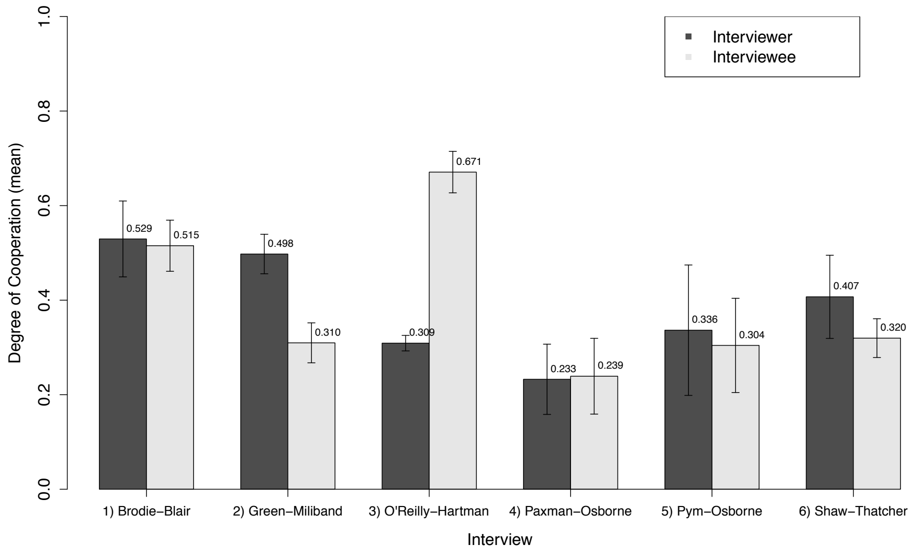


Figure 4.20: Degree of cooperation for the interviews in the annotated corpus (mean with error bars)

4.5.2 Eliciting Human Judgement on Cooperation

We obtained judgements on the behaviour of participants in the political interviews in the corpus by means of an online survey constructed using SurveyMonkey[32]. Observers were shown transcripts of the dialogues and asked to rate the behaviour of the participants with respect to their intuitions on how a political interview should normally go. Further details of the study are given below.

Materials

We used the six interviews in the corpus described above in Section 4.3.1. Judges were shown the same context and transcript as the annotators. The entire survey is available in Appendix C.

Participants

Volunteers were invited to take part in the study via email to the mailing lists of four research special interest group, a forum of postgraduate students in computing and via a series of posts on the social networking site Facebook[33].

Participants were not required to judge all the interviews in the survey. After responding to the mandatory questions about their cultural background and experience on dialogue analysis on the first page, they were given the option to skip pages or exit the survey. A total of 98 volunteers entered the questionnaire. Of these, 30 judged all 6 interviews and 24 provided judgements for at least one of them. The remaining 44 were discarded. Figure 4.21 summarises the English proficiency, cultural background

[32] http://www.surveymonkey.com
[33] The contents of the email and the Facebook post are reproduced at the end of Appendix C.

and dialogue analysis expertise for the 54 remaining volunteers[34]:

English proficiency: 23 (45.6%) native English speakers, 25 (46.3%) non-native fluent English speakers and 6 (11.1%) non-native with some knowledge of English.

Cultural background: 17 (31.5%) British or American, 11 (20.4%) non-British and non-American who have lived in the UK/US for more than five years, 5 (9.3%) non-British and non-American who have lived in the UK/US for less than five years, 4 (7.4%) non-British and non-Ammerican but Commonwealth and 17 (31.5%) from other backgrounds.

Dialogue analysis expertise: 4 (7.4%) experts, 21 (38.9%) with some experience in research, 16 (29.6%) with some informal experience and 13 (24.1%) with no experience whatsoever.

Design and Implementation

We designed the study as a survey with 16 questions grouped in 8 pages. The first page has the three mandatory background questions presented above. Pages 2 to 7 are one for each of the six interview fragments with the context, the dialogue transcripts and two questions: one asking for an assessment of the behaviour of the participants in the interview and one about the judge's familiarity with the dialogue and its context (see Figure 4.22). As stated above, replying to this questions was optional and volunteers could skip through any of them. Bearing this in mind and to increase coverage, the order of these pages was randomised. The last page has one optional question

[34] Age range and sex of the participants are missing here as they were not part of the information collected in the survey.

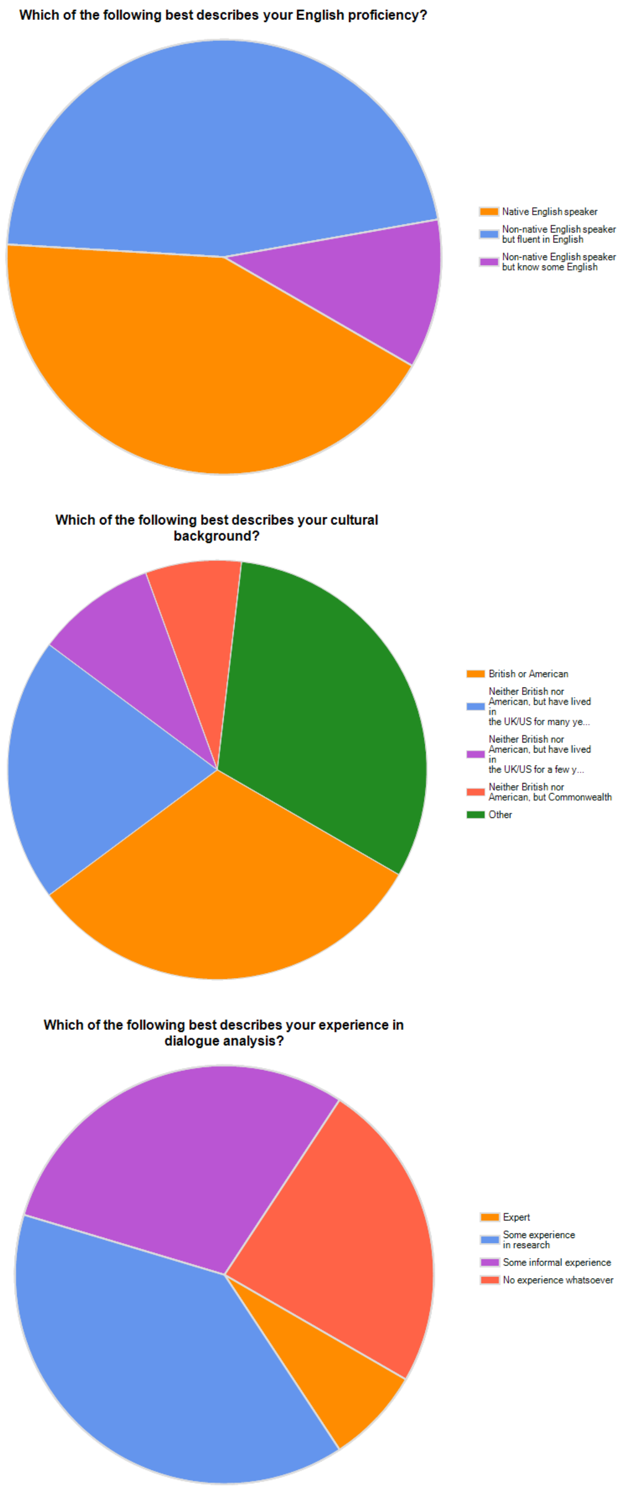


Figure 4.21: English proficiency, cultural background and dialogue analysis expertise of survey volunteers

asking volunteers whether they read carefully the interviews they judged and a text box for comments. Figure 4.22 shows the questions that were asked after the volunteers had read each of the interview transcripts, asking them to rate the performance of the participants based on their intuitions on how interviewers and politicians ought to behave. They were given five options, which ranged from Incorrect to Correct.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in this fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
Brodie	○	○	○	○	○
Blair	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

Next

Figure 4.22: Questions put to human observers for eliciting their judgement on the behaviour of the participants in a political interview

To implement the study we used SurveyMonkey[35]. We chose this tool as it provided a quick and easy way of laying out the materials, with little preparation of the data and minimal development. It also facilitated deployment, as the survey was accessible from any computer with internet access, as well as the collection, filtering and some analysis of the responses. Appendix C shows the exact wording and layout of the questions.

Results

The number of judgements received by the speakers in each interview are as follows:

[35] http://www.surveymonkey.com

1. Brodie (46) and Blair (44)
2. Green (35) and Miliband (36)
3. O'Reilly (39) and Hartman (39)
4. Paxman (42) and Osborne (40)
5. Pym (36) and Osborne (35)
6. Shaw (36) and Thatcher (36)

Figures 4.24 and 4.25 show the distribution of these judgements (left) with charts summarising the responses to the familiarity questions (right) for each survey. The aggregated judgements are shown on Table 4.23 and as a bar chart with error bars in Figure 4.26.

Of the 54 respondents, 31 replied to the question on the last page asking whether they had read the interviews they judged in detail: 25 (80.6%) answered positively and 6 (19.4%) negatively. There were 7 comments left in the text box, which are reproduced in Figure 4.23.

4.5.3 Correlation Analysis Between Survey Results and the Degree of Cooperation

We studied the relation between human judgement resulting from the survey and the degree of cooperation obtained from the method described above by means of a correlation analysis[36]. The outcomes of the method applied to the annotations produced by the coders were aggregated for the speakers in each interview fragment (cf. Table 4.22 and Figure 4.20). These values were plotted against the corresponding survey results also aggregated across judges for the speakers in each interviews (cf. Table 4.23 and Figure 4.26). This results in the 12 points presented as a scatter plot in Figure 4.27 showing error bars, the regression line and the value of Pearson's correlation coefficient $r = 0.471$. This value of Pearson's r is interpreted a moderate-to-strong positive correlation. However, its statistical significance is weak

Feedback/Criticism After watching the interviews (below) I got the impression that an analysis of the spoken interviews might have given a different analysis than the analysis of the written transcripts.
9/17/2012 12:20 AM View Responses Categorize As ▼

Feedback/Criticism I thought that the transcriptions, including comments such as (interrupting) may influence the reader to think that the behaviour of the speaker is not correct. Maybe it would have been better to watch the interviews rather than read a transcript?
9/4/2012 10:18 PM View Responses Categorize As ▼

Feedback/Criticism There's obviously a lot of context here; Paxman's interview with Osborne was fine because as a senior UK politician (I assume the interview's taken from Newsnight), Osborne and the audience know full well how the interview is going to be carried out. It would be quite inappropriate for, say, Huw Edwards to attempt a similar style on the News at Ten, or for Paxman to conduct a similarly aggressive interview with an academic statistician or representative of a small local pressure group (although he often does do this). So if I knew more about the the US interviewers/interviewees, I might have rated them differently. Of course, asking "Is this appropriate /for that interviewer/" is only possible for respondents who know the interviewer and interviewee.
9/3/2012 12:13 AM View Responses Categorize As ▼

Feedback/Criticism I read about 2/3rds of them in detail. I'd like to see how you detach sentiment bias from correct behaviour opinions! thanks for the extra interviews, too :)
8/31/2012 3:35 PM View Responses Categorize As ▼

Feedback/Criticism Very interesting, but I wondered what "ought to" meant in this context. Ought to according to whom? Rhetoricians? Politicians? The public?
8/31/2012 7:11 AM View Responses Categorize As ▼

Feedback/Criticism As an American, I'm not so exposed to day to day politics in the UK, but some of these clips are amazing. The first one was not far removed from the canonical, "When did you stop beating your wife?" heh Good luck with your research!
8/30/2012 11:23 PM View Responses Categorize As ▼

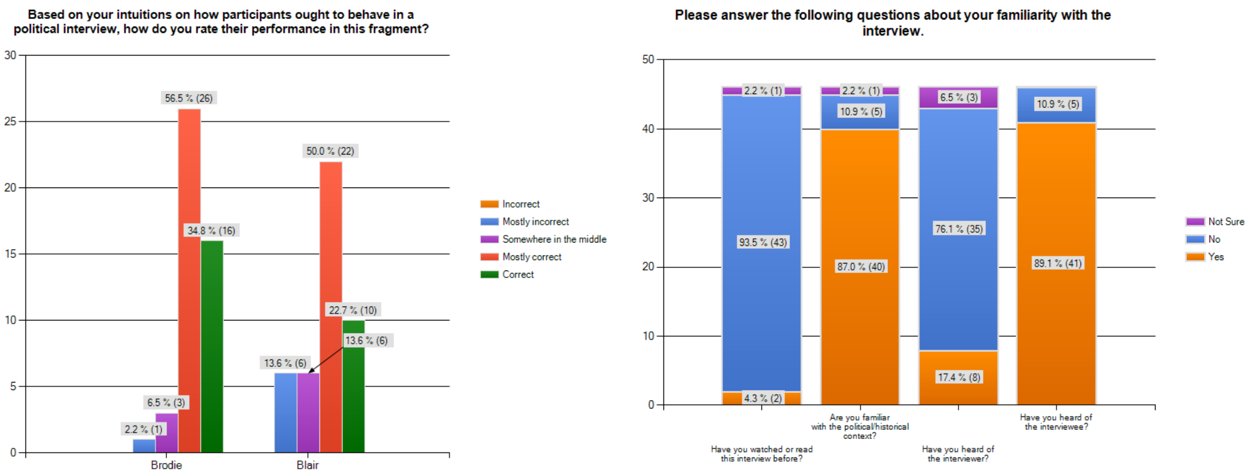
Feedback/Criticism The fragments are very long and it would require quite a lot of thought to analyse thoroughly. I was also a bit confused by the question around whether the participants behaved correctly based on my opinion of how people should behave in political interviews. This seems a bit vague to me and I've never given much thought to how participants should behave in a political interview versus any other interview. I based my responses on whether or not I felt the interviewee was answering the question posed or whether they were trying to avoid answering.
8/30/2012 7:27 PM View Responses Categorize As ▼

Figure 4.23: Feedback comments left by volunteers at the end of the survey

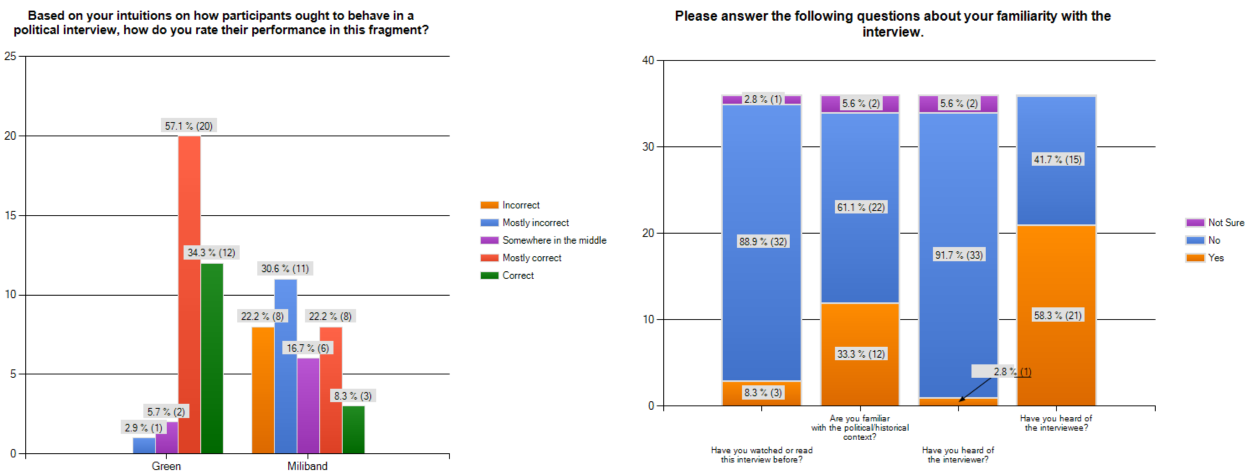
for that number of cases ($p = 0.123$).

For further insight, we carried out a similar analysis separating interviewers from interviewees. The rationale for this step is that some of the rules of the dialogue game are role-specific, making the method strictly different for each participant in an interview. A similar argument applies to the way human observers are expected to judge the behaviour of interviewers and politicians. The two sets of six points are shown in Figure 4.28, with separate regression lines and values for Pearson's r. The results show that correlation is significantly better for interviewers ($r = 0.753$) than for

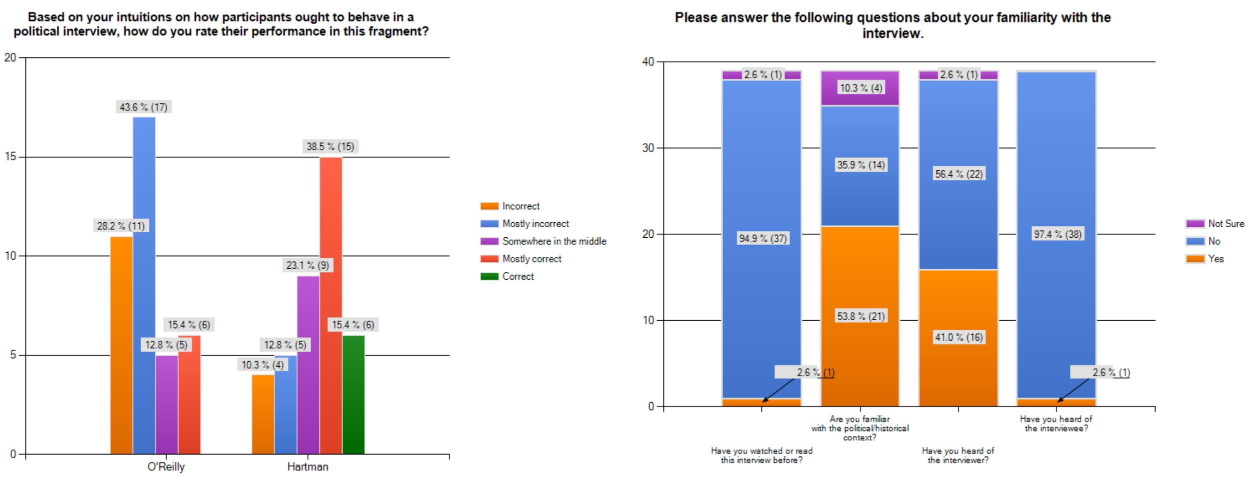
[36] The analysis of correlation and the plots were done using the R software environment for statistical computing and graphics (R Development Core Team, 2011).



(a) Interview 1: Brodie and Blair

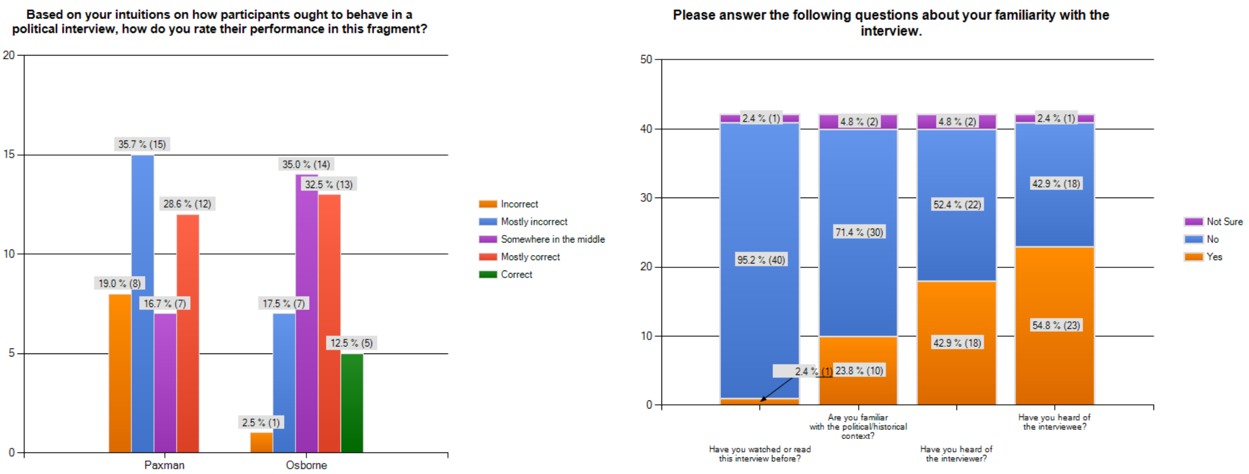


(b) Interview 2: Green and Miliband

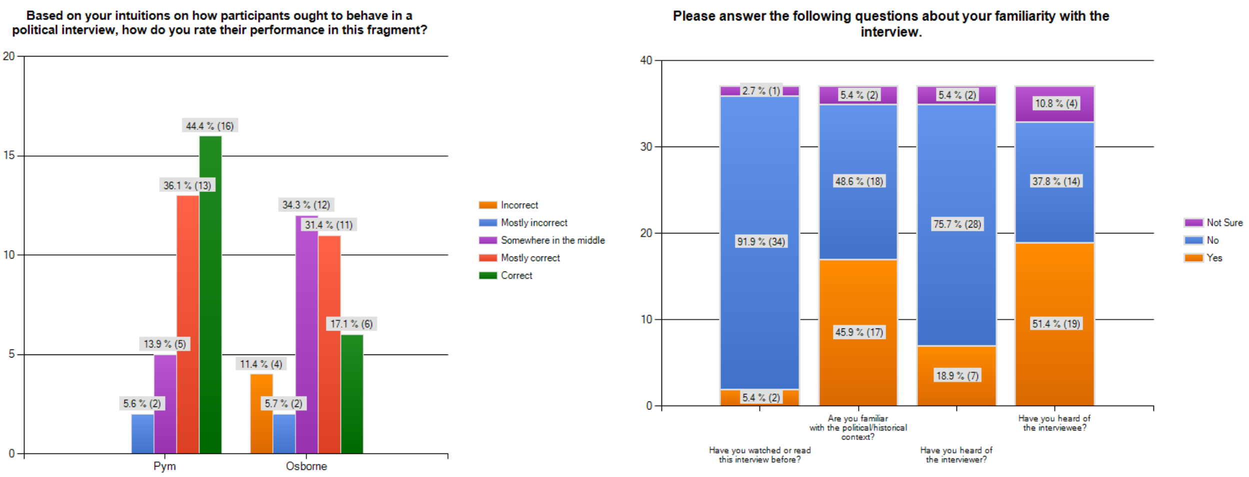


(c) Interview 3: O'Reilly and Hartman

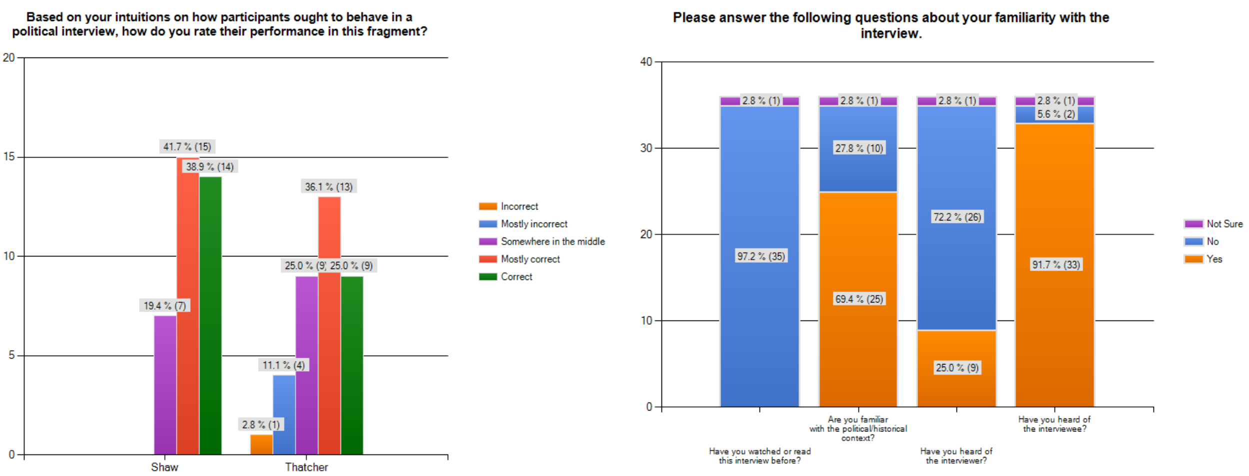
Figure 4.24: Human judgements on participant behaviour and familiarity of volunteers with the interview



(a) Interview 4: Paxman and Osborne



(b) Interview 5: Pym and Osborne



(c) Interview 6: Shaw and Thatcher

Figure 4.25: Human judgements on participant behaviour and familiarity of volunteers with the interview (continued)

Table 4.23: Human judgement on participant behaviour for the interviews in the corpus

Interview	Mean	Interview	Mean
1. Brodie and Blair		**4. Paxman and Osborne**	
INTERVIEWER	4.239	INTERVIEWER	2.548
INTERVIEWEE	3.818	INTERVIEWEE	3.350
2. Green and Miliband		**5. Pym and Osborne**	
INTERVIEWER	4.229	INTERVIEWER	4.194
INTERVIEWEE	2.639	INTERVIEWEE	3.371
3. O'Reilly and Hartman		**6. Shaw and Thatcher**	
INTERVIEWER	2.154	INTERVIEWER	4.194
INTERVIEWEE	3.359	INTERVIEWEE	3.694

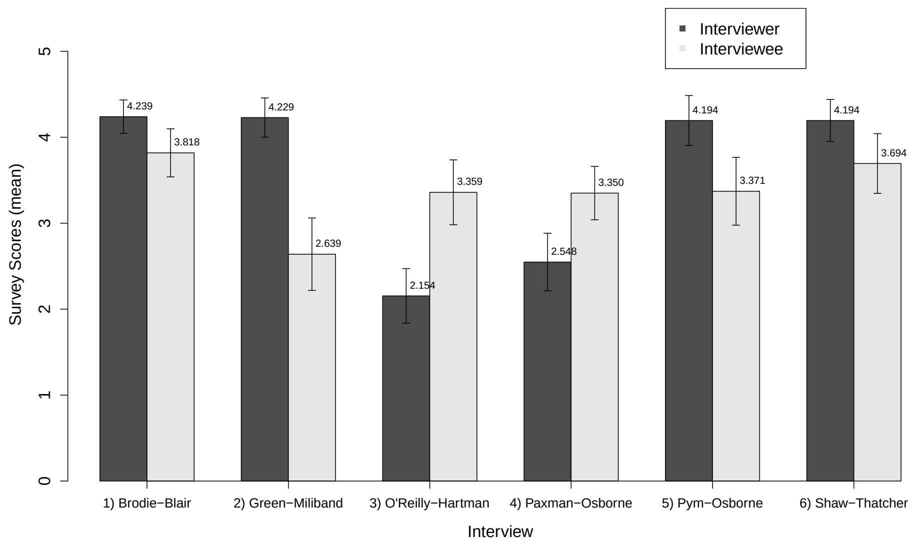


Figure 4.26: Human judgement on participant behaviour for the interviews in the corpus (mean with error bars)

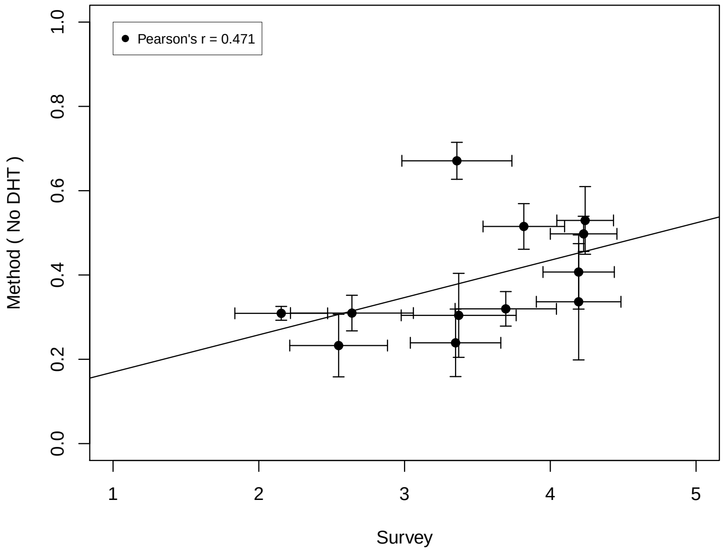


Figure 4.27: Survey results and the degree of cooperation for the political interviews in the corpus (means with error bars, regression line and Pearson's r correlation coefficient)

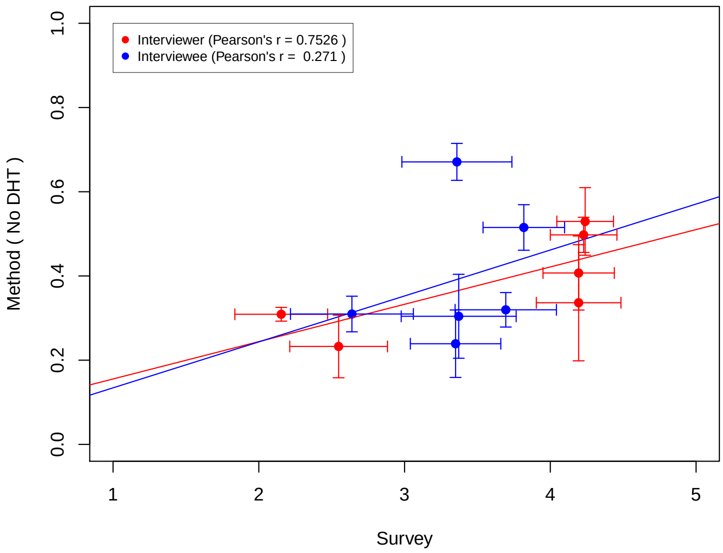


Figure 4.28: Survey results and the degree of cooperation for interviewers and interviewees (means with error bars, regression line and Pearson's r correlation coefficient)

interviewees ($r = 0.271$). Statistical significance is also stronger for interviewers with respect to the combined set ($p = 0.084$) indicating a trend towards positive correlation between the results of our method and human judgement. For the interviewees the correlation is not statistically significant ($p = 0.603$). With a sample of this size, correlation analysis is fairly sensitive to outliers, which could explain such a high p-value for the interviewees. Take, for instance, the behaviour of the interviewee in Interview 3 (O'Reilly and Hartman), which scored fairly high via the method (0.671/1.000) but just above average on the survey (3.359/5.000). This corresponds to the point furthest up from the regression line for interviewees in Figure 4.28. Coincidentally, Interview 3 has been described by one of the annotators as "more like a debate than a political interview" which could explain the unexpected value given by the method. The values of Pearson's r and p are shown for each role and for the combined set are on Table 4.24.

Table 4.24: Pearson's correlation coefficient between survey results and the degree of cooperation for the political interviews in the corpus (both aggregated)

	Pearson's r	**p-value**
Dialogue	0.471	0.123
Interviewer	0.753	0.084
Interviewee	0.271	0.603

Dialogue History Threshold

One criticism of the degree of cooperation computed as described earlier in the chapter is that it can penalise participants too harshly for obligations introduced early in the dialogue and never met. As these stay in the set of pending obligations, they are considered dynamic non-cooperative features and contribute to the amount of non-cooperation each time the participant releases the floor without discharging them. This might be unrealistic from the point of view of human judgement, especially for obligations like the rejection of invalid contributions. For this reason, we introduced a runtime parameter called **dialogue history threshold** (DHT) which controls the number of turns after which unmet obligations cease to be considered non-cooperative features.

We ran the method on the same annotated data for 5 values of the DHT parameter. The aggregated results for the participants in each interview are presented in Table 4.25 and shown as bar charts with error bars in Figures 4.29 and 4.30. As expected, the degree of cooperation increases with lower values of the DHT as a larger number of non-cooperative features are ignored.

Table 4.25: Degree of cooperation for the interviews in the annotated corpus for different values of the dialogue history threshold (DHT) runtime parameter (means)

Interview	No DHT	DHT=5	DHT=4	DHT=3	DHT=2	DHT=1
1. Brodie and Blair						
Interviewer	0.529	0.678	0.681	0.794	0.798	0.986
Interviewee	0.515	0.602	0.640	0.644	0.691	0.695
2. Green and Miliband						
Interviewer	0.498	0.537	0.557	0.614	0.64	0.801
Interviewee	0.310	0.341	0.364	0.382	0.44	0.481
3. O'Reilly and Hartman						
Interviewer	0.309	0.412	0.449	0.464	0.530	0.593
Interviewee	0.671	0.694	0.694	0.741	0.741	0.850
4. Paxman and Osborne						
Interviewer	0.233	0.325	0.360	0.394	0.468	0.530
Interviewee	0.239	0.330	0.349	0.403	0.448	0.512
5. Pym and Osborne						
Interviewer	0.336	0.426	0.431	0.534	0.544	0.800
Interviewee	0.304	0.344	0.371	0.389	0.439	0.468
6. Shaw and Thatcher						
Interviewer	0.407	0.533	0.568	0.594	0.638	0.740
Interviewee	0.320	0.485	0.525	0.572	0.629	0.713

Figures 4.31 and 4.32 show the scatter plots for the values of the DHT parameter from 5 to 1, both inclusive. The plots on the left are for the combined sets of judgements and on the right they are segregated per speaker

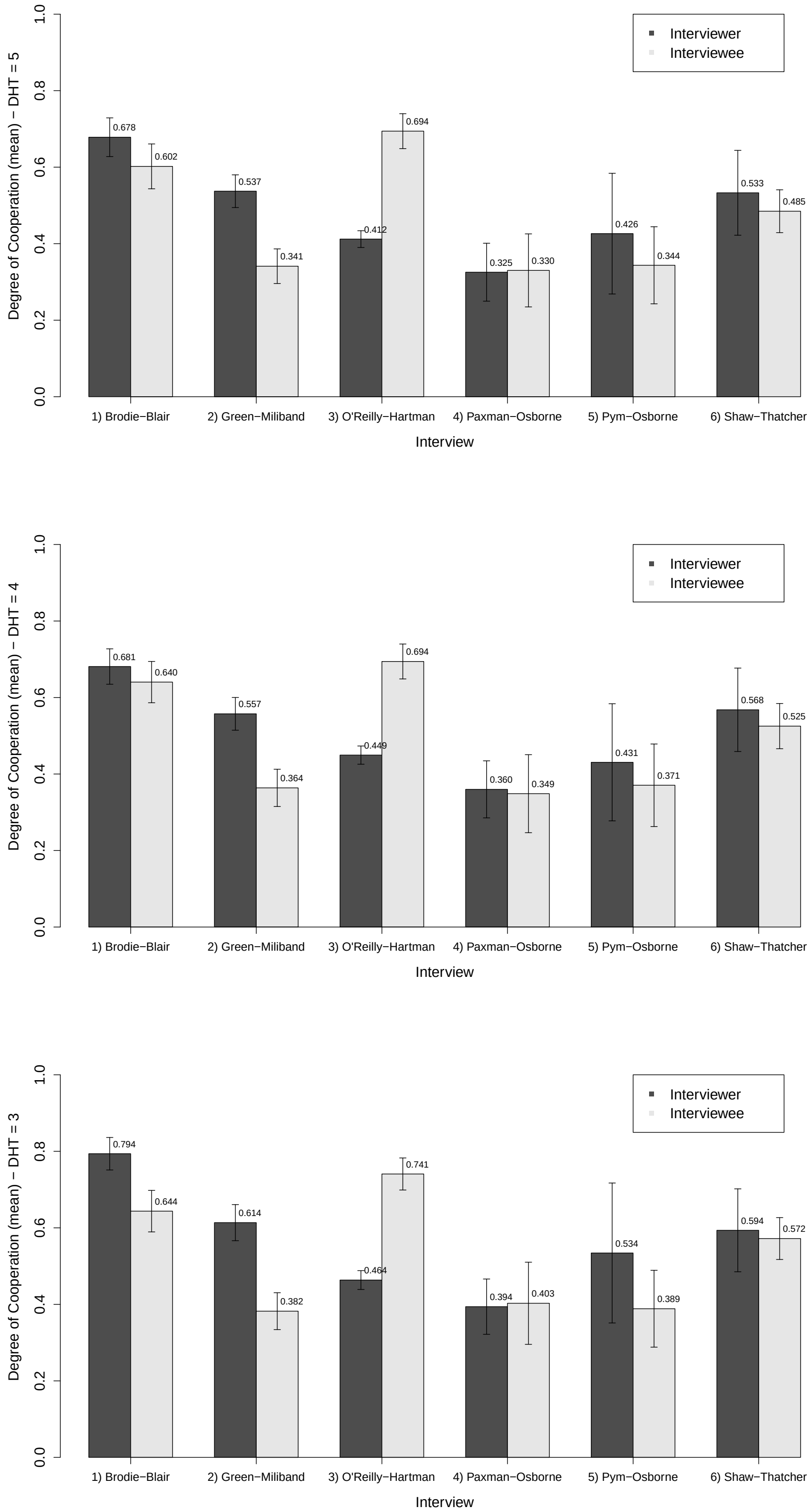
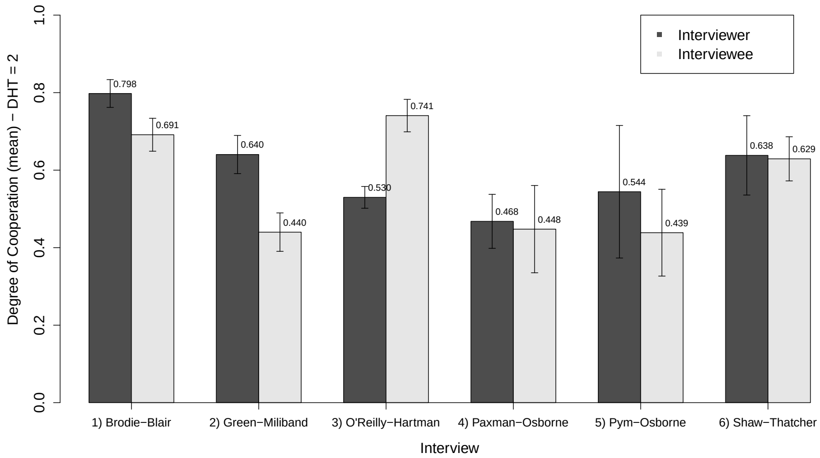


Figure 4.29: Degree of cooperation for the interviews in the annotated corpus for different values of the dialogue history threshold (DHT) runtime parameter (mean with error bars)



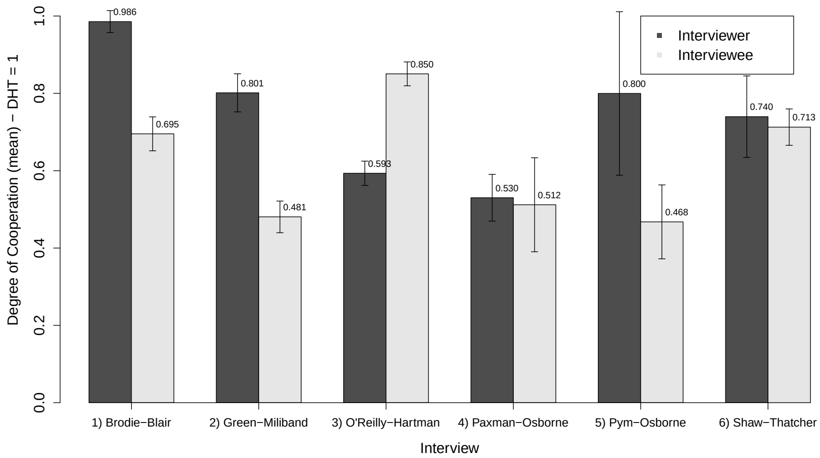
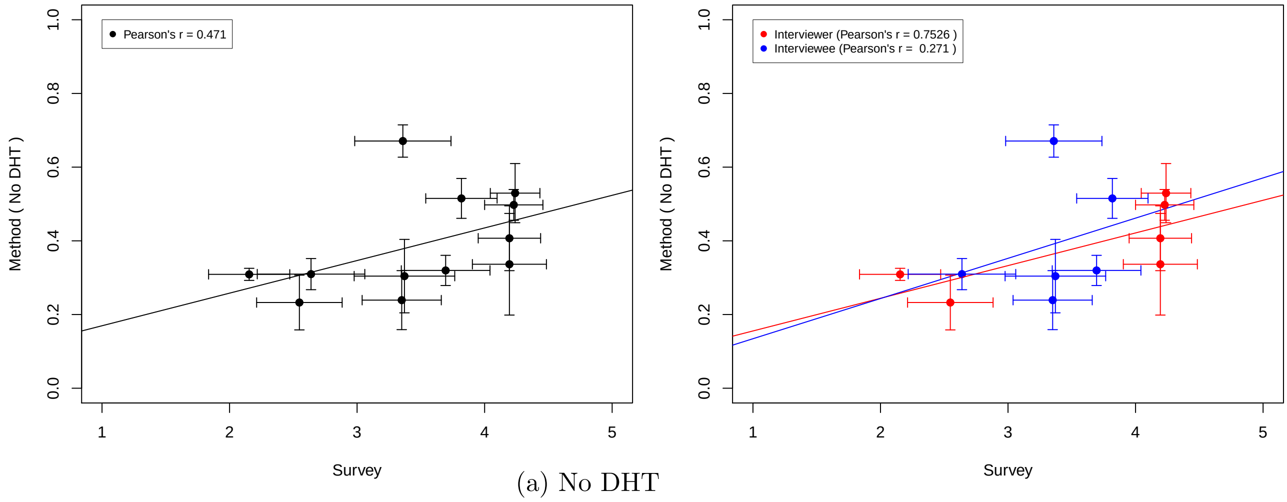


Figure 4.30: Degree of cooperation for the interviews in the annotated corpus for different values of the dialogue history threshold (DHT) runtime parameter (mean with error bars; continued)

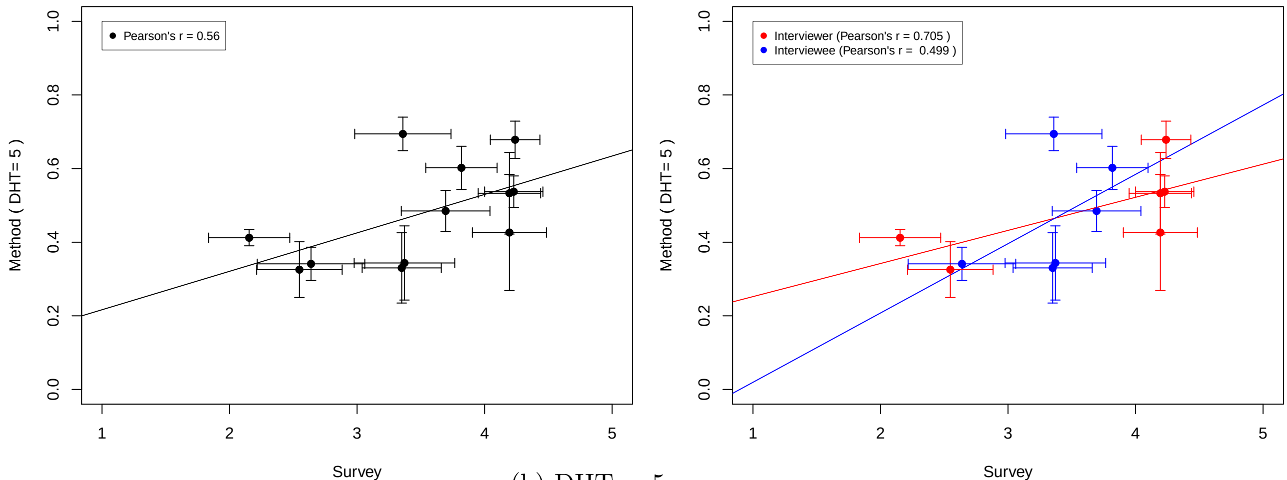
role, as presented earlier for the results without dialogue history threshold[37]. The analysis of correlation improved considerably with the introduction of the parameter. For the combined set of observations, the three lower values of the threshold give statistically significant results ($p = 0.028, 0.044$ and 0.012, for DHT $= 3, 2$ and 1, respectively) with strong positive correlation (Pearson's $r = 0.631, 0.589$ and 0.697, respectively). The other two values for the parameter, 4 and 5, also indicate a trend towards strong positive correlation ($r = 0.560$ and 0.542, respectively, with $p = 0.058$ and 0.069). These values are shown in Table 4.26 on the first row.

The analysis of correlation for each conversational role reveals issues similar to the results before the introduction of the DHT parameter. Correlation is substantially stronger – and statistically more significant – for interviewers than it is for interviewees. This seems to indicate that the issues with the latter set that we discussed above persist after introducing the parameter. In fact, as can be seen in Table 4.26, correlation for interviewers decreases with the introduction of the threshold and only improves for DHT $= 1$. On the other hand, correlation for interviewees increases significantly with all the values of the parameter, peaking for DHT $= 2$. Overall, the best results seem to be given by a dialogue history threshold of 1. This means that only obligations introduced in the same turn that are not discharged contribute to the degree of cooperation as perceived by the human observers. Such obligations emerge, for instance, when interviewers implicitly or explicitly accept a question and then do not provide a valid reply, or when interviewers fail to ask a valid question.

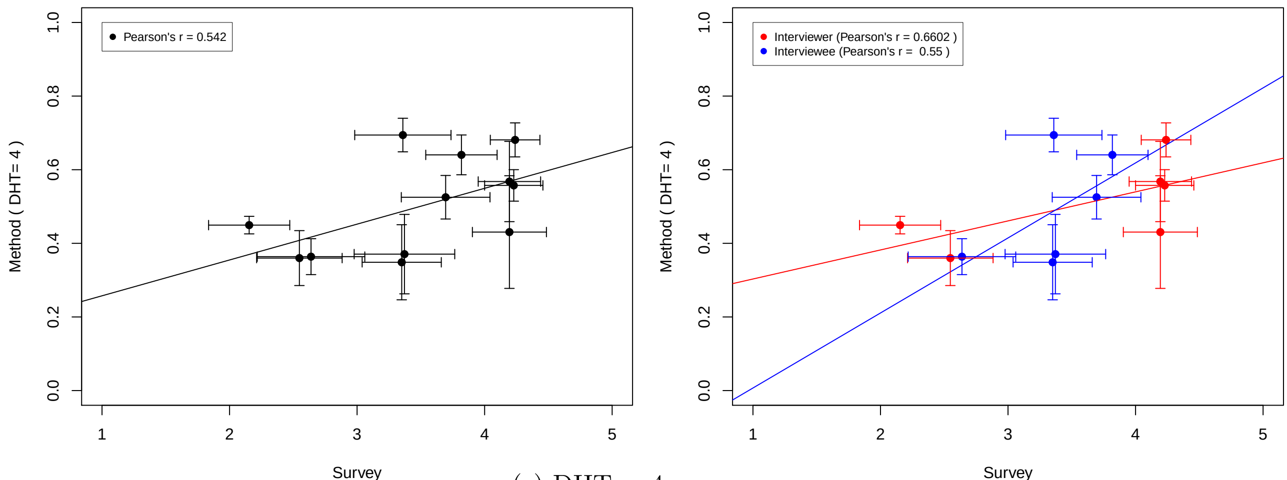
[37] To aid the comparison, we repeat here the results shown above in Figures 4.27 and 4.28



(a) No DHT

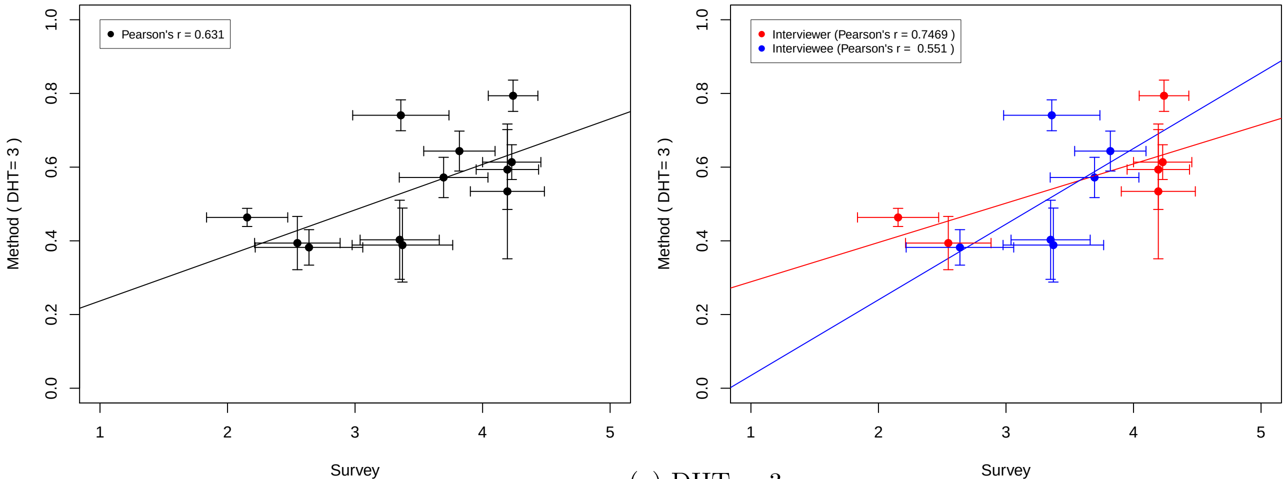


(b) DHT = 5

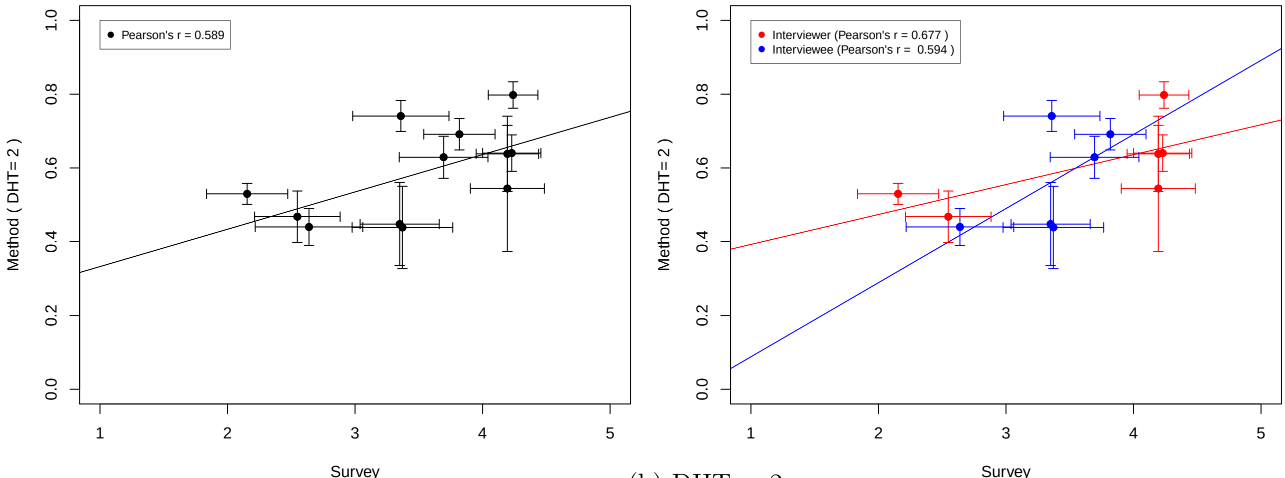


(c) DHT = 4

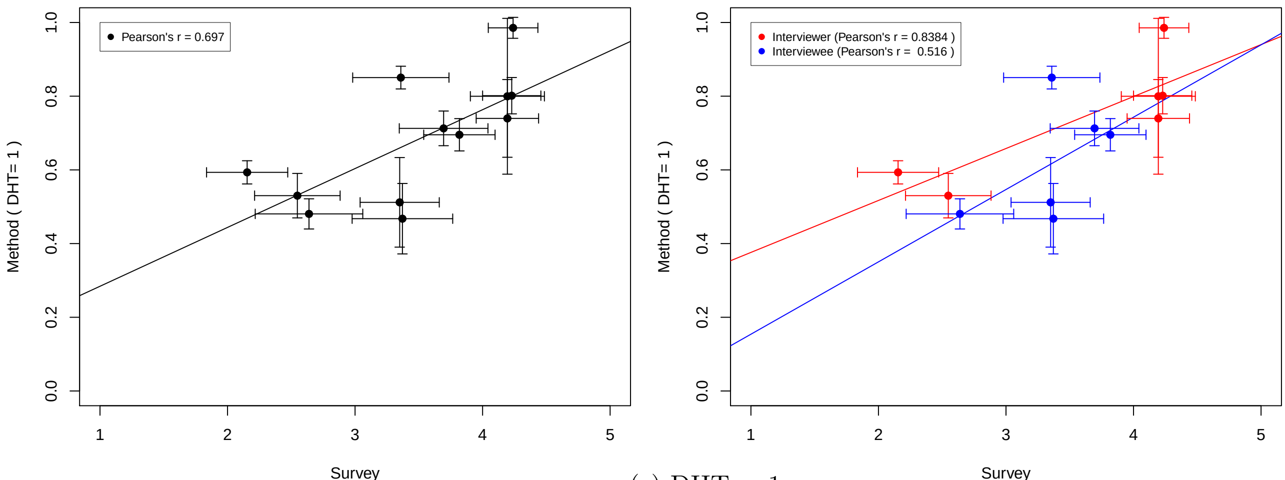
Figure 4.31: Correlation between survey results and the degree of cooperation for different values of the dialogue history threshold (DHT) runtime parameter



(a) DHT = 3



(b) DHT = 2



(c) DHT = 1

Figure 4.32: Correlation between survey results and the degree of cooperation for different values of the dialogue history threshold (DHT) runtime parameter (continued)

Table 4.26: Pearson's correlation coefficient between survey results and the degree of cooperation for different values of the dialogue history threshold (DHT) runtime parameter

	No DHT		**DHT = 5**		**DHT = 4**		**DHT = 3**		**DHT = 2**		**DHT = 1**	
	r	**p**	**r**	**p**	**r**	**p**	**r**	**p**	**r**	**p**	**r**	**p**
Dialogue	0.471	0.123	0.560	0.058	0.542	0.069	0.631	0.028	0.589	0.044	0.697	0.012
Interviewer	0.753	0.084	0.705	0.118	0.660	0.154	0.747	0.088	0.677	0.141	0.838	0.037
Interviewee	0.271	0.603	0.499	0.314	0.550	0.259	0.551	0.257	0.594	0.214	0.516	0.295

4.6 Discussion

The method presented above is, to date and to the extent of our knowledge, the most elaborate attempt at annotating and analysing naturally occurring dialogue in the light of linguistic cooperation. Also novel is the application of such an approach to a corpus of real political interviews, especially in that both speakers received the same amount of attention and that the method was subject to an extensive evaluation.

The results of the evaluation for reliability are encouraging and indicate that the method is suitable for the systematic analysis of non-cooperation. They also expose some of its weaknesses, such as the difficulties for applying some of the criteria in the manual annotation, a degree of vagueness in the definition of a few of the concepts and the inherent subjectivity of many of the judgements involved in properly characterising non-cooperation. A revision of the guidelines as discussed in Section 4.3 would contribute in this direction.

The evaluation of validity produced fairly good results, especially considering how little information was given to observers in the survey as to what was meant by linguistic cooperation and the total absence of a reference to the specific dialogue game adopted as part of the semi-automatic measure.

It is worth pointing out that the method, in its current form, was able to predict accurately in the six interviews of the corpus the participants that behaved better with respect to their interlocutors. Beyond the correlation of the precise scores, the ability to determine this binary judgement without mistakes in all cases is of great interest and an indication of the adequacy of the approach.

It is unfortunate that the size of the sample in the corpus prevented from obtaining statistically significant results for each speaker role, particularly the interviewee. Strictly speaking, as the rules that are used to assess how cooperative participants are depend on their specific roles, the method is in fact distinct for each role. A larger sample, including more interview fragments would help in setting this right. Given the relative ease in collecting human judgements, the inclusion of new fragments should start with one or more surveys similar to the one described above. This would allow to decide what subset of interviews offers the best coverage of the range of possible behaviours – something that was somewhat missing in our study. Only once this set has been chosen, it would commence the annotation of the transcripts, a procedure considerably more expensive in terms of the resources it involves.

One key element of the method is the dialogue game described in Chapter 3. As it specifies the rules against which the behaviour of the participants in a dialogue is assessed, the accuracy of the measure of cooperation depends closely on the dialogue game. A revision of these rules to draw from the intuitions human observers use when assessing linguistic cooperation could result in a more accurate measure and an increase in the correlation between the results of the measure and human judgement. Another possibility – or perhaps a complement – would be to extract the rules automatically from a corpus annotated with the scheme described in this chapter. We will come back to this in Chapter 6.

4.7 Summary

In this chapter we described a semi-automatic empirical method for measuring linguistic non-cooperation in naturally occurring dialogue. The method, instantiated for the analysis of political interviews, was applied to a corpus of six interview fragments. The evaluation of the method consisted of two parts: a detailed inter-annotator agreement analysis on the annotated corpus for assessing reliability and a correlation analysis study between the results of the method and human judgement elicited by means of an online survey.

The method builds on the concepts introduced in the previous chapter and bridges the gap between those concepts and phenomena that are present in naturally-occurring dialogue. By analysing inter-annotator agreement, we evaluated the reliably of this connection. By analysing the correlation of the results of the method with human judgement, we assessed the validity of the link between the conceptual framework (e.g. dialogue game rules, discourse obligations and the proposition that linguistic cooperation can be regarded as the extent to which dialogue participants behave withing the rules of the game) and the way human observers perceive linguistic cooperation in dialogue by means of the empirical measure.

In the next chapter we propose and discuss an architecture for conversational agents that incorporates the notions introduced in Chapter 3 and the automatic aspects of the method discussed above to interpret, reason and act in conversation.

Chapter 5

Modelling Non-Cooperative Conversational Agents

This chapter describes a modelling approach for constructing conversational agents that can exhibit and deal with diverse degrees of linguistic cooperation as defined and analysed in previous chapters. Among other elements, the model includes: a description of the dialogue game as defined in Chapter 3, a specification of the conversational domain in which semantic relations such as the validity of questions and replies are specified; a mechanism for assessing participant behaviour along the lines of the method described in Chapter 4 and an agenda of private goals.

Section 5.1 describes the elements above in detail and how they relate to each other and to the analysis in the previous chapter. Section 5.2 shows how the model combines these elements into agents that can interact producing dialogues that resemble the example first presented in Figure 3.6 in Chapter 3 and reproduced here in Figure 5.1, but that can range in the level of cooperation displayed by the agents depending on very simple parameter settings. We compare our model to related approaches in Section 5.3 and refer to a proof-of-concept, prototype implementation that informed the development of many of our ideas in Section 5.4.

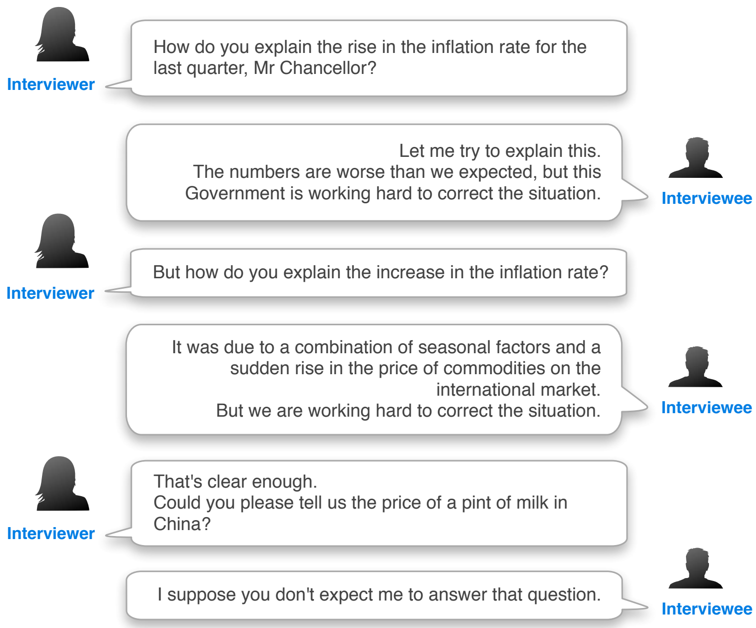


Figure 5.1: A (hand-crafted) political interview example (D_1)

5.1 Non-Cooperative Conversational Agents

Conversational agents must deal with several informational and interactional aspects throughout a dialogue. For instance, they need knowledge about what conversational actions are available to them, about what goals they want to achieve in the conversation and about what constraints are imposed upon them by the type of dialogue in which they are engaged. They also need to keep track of the dialogue dynamics, that is the dialogue history, the goals they have achieved and those yet to be pursued and any pending discourse obligations on either side. For the model we propose in this chapter, we group these elements in separate modules as follows:

1. A taxonomy of **dialogue acts** specifying the conversational actions that participants perform and understand in the dialogue, and a taxonomy of **action labels**, as described in previous chapters.

2. An **information state** to keep track of the dialogue history, dynamic obligations and private goals.

3. Mechanisms for **understanding** and **generating natural language** utterances.

4. A specification of the agent's **role**. This is the management of aspects specific to each agent, such as the identity of the role they play in the dialogue, the agenda of private goals and so forth.

5. A **conversational domain** defining valid and invalid contributions, relative relevance, e.g. of answers to questions, etc.

6. A **dialogue game** with the rules specifying the actions allowed to the participants' roles and the dynamics of discourse obligations.

7. A **deliberation mechanism** to decide on their next contribution.

8. A **control algorithm** that binds these elements together and interacts with the other agent by writing and reading utterances to and from shared channels.

These elements are shown in Figure 5.2. The modules with dashed borders allow modelling different types of dialogue, roles and conversational domains. They are instantiated according to specific aspects of the type of dialogue being modelled as shown in Figure 5.3.

In the rest of the section, we formalise these elements and illustrate with examples from the political interview dialogue setting. Before, let us make a few modelling assumptions with the aim of scoping the discussion:

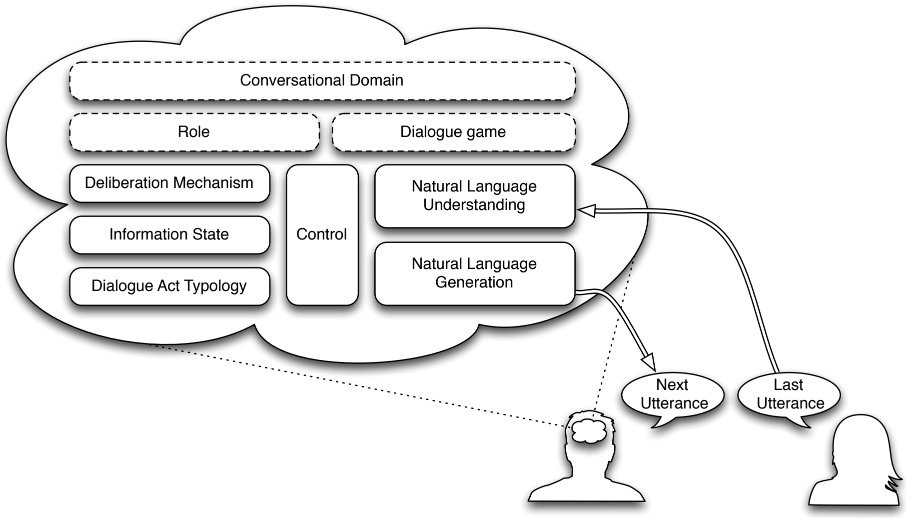


Figure 5.2: Elements in a conversational agent

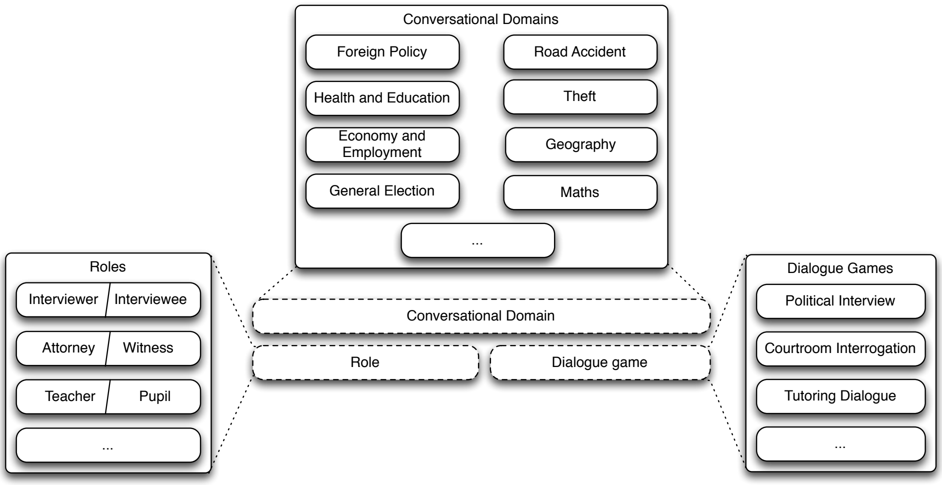


Figure 5.3: Domain-specific elements in a conversational agent

- Agents are autonomous and do not physically share any of the modules above. However, we assume that they agree on the dialogue act and action label taxonomies, on the structure of the information state, on the conversational domain and on the dialogue game. In fact, all these modules are assumed to be identical in both agents, with the exception of the agent's role, the control algorithm and the information state.

- Language understanding and generation are assumed to be flawless. This means that the agents always agree on what is being communicated in either direction, relieving us from having to deal with management of grounding, evidence of understanding, and so forth.

- We consider differences in the contents of the information state, in particular for the agenda of private goals. Each agent keeps a specification of its private goals and of the management of the agenda, but not that of the other party.

- The deliberation mechanisms can also differ, implementing distinct criteria for selecting an agent's next contributions and to reflect which party holds the floor at the start of the conversation.

We will discuss the possibilities for future work that result from relaxing the assumptions above in Chapter 6.

5.1.1 Dialogue Act and Action Label Taxonomies

Dialogue acts are the basic conversational actions that agents can perform and understand. For generality, we model dialogue acts as having **function** and **content**. Formally, we represent dialogue acts as $F(c)$, where F is the function and c the content. The performance of a dialogue act is called a **dialogue move** and is represented as a pair $m = (s, F(c))$ where s is the speaker performing the dialogue act $F(c)$.

As an example, for political interviews we consider the following dialogue acts:

- $Question(p)$: an information request. The content is some representation of the propositional content of an utterance.

- $Statement(p)$: an assertion providing information. The content is also some representation of the propositional content of an utterance.

- $Accept(m)$: the acceptance of a previous contribution. The content is a dialogue move.

- $Reject(m)$: the rejection of a previous contribution. The content is also a dialogue move.

We leave the representation for the propositional content of an utterance unspecified. They can range from simply the surface realisations produced by the speaker to deep semantic representations (Traum, 2003). For the example we use unique identifiers, understood as pointers to the surface utterances (cf. Section 5.1.3). Hence, a question is represented as $Question(q_1)$ and an acceptance of a move in which this question is asked by speaker s is represented as $Accept((s, Question(q_1)))$.

Action labels were introduced and formalised in Chapter 3 (page 75) for representing the actions of dialogue participants in terms that were compatible with the specification of the rules of the dialogue game. The taxonomy of action labels for political interview agents was described in Section 3.5.1 and used in Chapter 4 for the automatic assessment of cooperation in annotated dialogue:

$$(\langle \mathit{ID} \rangle) : \langle \mathit{action\ name} \rangle [@(\langle \mathit{referent\ ID} \rangle)] \ \ [\langle \mathit{repeated\ \&\ complete\ flags} \rangle]$$

where the action name is in the set $T_{PI} = \{$`valid-statement`, `invalid-statement`, `valid-question`, `invalid-question`, `valid-reply`, `invalid-reply`, `acceptance`, `rejection`$\}$. Recall that only responsive actions – that is replies, acceptances and rejections – include a referent identifier, to indicate the action they respond to. Also, valid and invalid questions have **New/Repeated** associated flags to indicate whether the question had been asked earlier in the dialogue. Valid and invalid replies have **Complete/Incomplete** associated flag, indicating whether the entirety of the information requested in a question is provided by the reply, including any preceding partial replies.

5.1.2 Information State

Agents use a structure called **information state** (Cooper and Larsson, 1999; Traum and Larsson, 2003) to keep track of the dynamic state of the dialogue at all times. The information state includes the history of dialogue moves performed by both participants, the state of the agent's agenda of private goals, the dynamic obligations of both parties, and possibly other elements specific to the type of dialogue the agents are having. We represent this structure as a **typed record** (Cooper, 1998). The type of a record is defined by an unordered list of named fields, each one with an associated type. These types can be another record type, a basic type or a type constructed on the basis of other types. Types can be constructed as the product of two or more types or by using type constructors, such as **seq** and **set** for sequences and sets, respectively. The fields in instances of a type record are accessed with the dot notation.

Our basic types are *Utterance*, for the speakers' utterances; *Act*, for the dialogue acts in the taxonomy; *Speaker*, for the speakers in the conversation; and *Label*, for the action labels introduced in Chapter 3. The type for dialogue moves is defined as $Move = \mathit{Speaker} \times \mathit{Act}$. The type for obligations is $\mathit{Obligation} = \mathit{Speaker} \times \mathit{Label}$. The type for information states is then defined as:

$$
\mathit{InfoState} = \begin{bmatrix} \textsc{DialogueHistory} & : & \textbf{seq } \mathit{Move} \\ \textsc{PrivateAgenda} & : & \textbf{set } \mathit{Act} \\ \textsc{Obligations} & : & \textbf{seq } \mathit{Obligation} \\ & (\ldots) & \end{bmatrix}
$$

These are typed records with at least three fields:

- The dialogue history (DialogueHistory), represented as a sequence of moves, at any time in the dialogue holds the sequence of moves performed so far by either speaker.
- The agenda of private goals (PrivateAgenda), represented as a set of acts that the agent aims at performing during the conversation[1].
- The agents' dynamic obligations (Obligations), represented as a sequence of obligations as discussed in Chapter 4.

We assume that together with the specification of the information state structure the module includes a constant – or 0-ary function – providing the empty information state:

$$
\mathit{initialise\text{-}state} \quad : \quad \mathit{InfoState}
$$

In the political interview setting, the information state has two extra elements: the question being currently discussed (CurrentQuestion), and a

[1] This is a simplification: as part of their private agendas, agents only consider acts that they themselves can perform – therefore these are dialogue acts instead of dialogue moves. A more realistic model would also allow for acts performed by others as part of an agent's private goals or, more generally, for states of the world represented as partially specified information states. A planning mechanism using the dialogue game and the conversational domain to reason about the agent's own actions that would lead the other party to performing the desired goals or to the desired state of the world would be required in such cases (Fikes and Nilsson, 1972; Zinn et al., 2002; Steedman and Petrick, 2007). For instance, if an interviewer wants a question to be answered by the interviewee, he or she would have to ask it in the first place.

set of proposed questions or answers (joint projects) that have not yet been accepted or rejected (PROPOSALS), paired with their respective acceptance or rejection obligations according to the dialogue game. The type for proposals is $Proposal = Obligation \times Move$. The dialogue state for agents in political interviews is as follows:

$$InfoState = \begin{bmatrix} \text{DIALOGUEHISTORY} & : & \textbf{seq}\ Move \\ \text{PRIVATEAGENDA} & : & \textbf{set}\ Act \\ \text{OBLIGATIONS} & : & \textbf{seq}\ Obligation \\ \text{PROPOSALS} & : & \textbf{set}\ Proposal \\ \text{CURRENTQUESTION} & : & Act \end{bmatrix}$$

with the following empty information state[2]:

$$initialise\text{-}state(\,) = \begin{bmatrix} \text{DIALOGUEHISTORY} & = & \langle\rangle \\ \text{PRIVATEAGENDA} & = & \{\} \\ \text{OBLIGATIONS} & = & \langle\rangle \\ \text{PROPOSALS} & = & \{\} \\ \text{CURRENTQUESTION} & = & None \end{bmatrix}$$

Of these fields, only the contents of the private agenda are unique to each agent. The contents of the rest of the information state, although not publicly accessible to the other party, will be identical in both agents following the assumption that natural language generation and understanding are flawless and that the agents agree on the specifications of the dialogue game and the conversational domain.

5.1.3 Natural Language Understanding and Generation

Agents use the **natural language understanding** and **generation** modules to turn utterances into dialogue moves and vice-versa. The modules

[2] The value *None* used to initialise current questions is a special element of all types that denotes the null value. In this case, it means that no question is currently under discussion.

provide two functions, respectively:

$$\begin{aligned} interpret &: Utterance \rightarrow Move \\ generate &: Move \rightarrow Utterance \end{aligned}$$

For the general model, the internal details of these functions are left unspecified. In practice, they could be implemented using many of the approaches to natural language understanding and generation described in the literature and in current research (Allen, 1995; DeVault et al., 2011; Reiter and Dale, 2000; DeVault et al., 2008).

As mentioned above, in the example we assume that generation and understanding are flawless. This means that $interpret(generate(m)) = m$, for any dialogue move m. In other words, a dialogue move realised with the natural language generator in one of the agents is interpreted as such by the natural language understanding module in the other agent. Furthermore, we will use canned messages to present the resulting dialogues and assume that these functions convert between strings and dialogue moves using unique identifiers for each string. Table 5.1 shows the conversion of the utterances from the example in Figure 5.1, plus a few more we will use later in the chapter. In this case, the functions *interpret* and *generate* simply implement this mapping without any further processing.

5.1.4 Role

The **Role** module deals with elements specific to each agent throughout the dialogue. These include the agent's identity, the management of private goals, the initial contents of the private agenda, and possibly further aspects depending on the conversational setting or knowledge about the world that

Table 5.1: Mapping between speaker utterances and dialogue moves.

Speaker	**Utterance**	**Dialogue Move**
INTERVIEWER	"How do you explain the rise in the inflation rate for the last quarter, Mr Chancellor?"	$(\mathtt{ir}, \mathit{Question}(q_1))$
INTERVIEWEE	"Let me try to explain this."	$(\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1))))$
INTERVIEWEE	"The numbers are worse than we expected, but this Government is working hard to correct the situation."	$(\mathtt{ie}, \mathit{Statement}(s_1))$
INTERVIEWER	"That's not answering the question."	$(\mathtt{ir}, \mathit{Reject}((\mathtt{ie}, \mathit{Statement}(s_1))))$
INTERVIEWEE	"It was due to a combination of seasonal factors and a sudden rise in the price of commodities on the international market."	$(\mathtt{ie}, \mathit{Statement}(s_2))$
INTERVIEWEE	"But we are working hard to correct the situation."	$(\mathtt{ie}, \mathit{Statement}(s_3))$
INTERVIEWER	"That's clear enough."	$(\mathtt{ir}, \mathit{Accept}((\mathtt{ie}, \mathit{Statement}(s_2))))$
INTERVIEWER	"Could you please tell us the price of a pint of milk in China?"	$(\mathtt{ir}, \mathit{Question}(q_2))$
INTERVIEWEE	"I suppose you don't expect me to answer that question."	$(\mathtt{ie}, \mathit{Reject}((\mathtt{ir}, \mathit{Question}(q_2))))$
INTERVIEWER	"Can you comment on the position of your office regarding the inflation rate? "	$(\mathtt{ir}, \mathit{Question}(q_3))$
INTERVIEWEE	"Right."	$(\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_3))))$
INTERVIEWER	"Thank you."	$(\mathtt{ir}, \mathit{Accept}((\mathtt{ie}, \mathit{Statement}(s_1))))$

is known to only one of the agents. The module provides four functions:

$$\begin{array}{lcl} \textit{identity} & : & \textit{Speaker} \\ \textit{initialise-agenda} & : & \textbf{set } \textit{Act} \\ \textit{select-goal} & : & \textit{InfoState} \rightarrow \textit{Act} \\ \textit{update-agenda} & : & \textit{InfoState} \times \textit{Move} \rightarrow \textbf{set } \textit{Act} \end{array}$$

Once again, for the general model the internal details of these functions are left unspecified. In the case of political interviews, the initialisation sets up the private agenda with the private goals of the agent. For interviewers, the initial agenda contains the set of questions they aim at asking in the dialogue. For example[3]:

$$\textit{initialise-agenda}_{\texttt{ir}}(\,) = \{\textit{Question}(q_2)\}$$

For interviewees, the private agenda has a set of statements they aim at including before the conversation is over. These messages could include, for instance, key points, party policy or a personal position with respect to current affairs. For example:

$$\textit{initialise-agenda}_{\texttt{ie}}(\,) = \{\textit{Statement}(s_1), \textit{Statement}(s_3)\}$$

These initial values mean that the interviewer aims at asking the question "*Could you please tell us the price of a pint of milk in China?*" at some point in the dialogue. Similarly, the interviewee intends to deliver the messages "*The numbers are worse than we expected, but this Government is working hard to correct the situation.*" and "*But we are working hard to correct the situation.*".

Agents might also want to reject or accept specific contributions from the other party, regardless of whether they are valid or not. In such cases,

[3] For the surface realisation of $\textit{Question}(q_2)$, see Table 5.1.

their agendas would contain accept and reject dialogue acts, with the moves they refer to as content. For an example, if in the dialogue of Figure 5.1 the interviewee had intended to reject the first question posed by the interviewer regardless of its validity, the act $Reject(Question(q_1))$ would be part of the interviewee's agenda of private goals.

It is worth noting that, contrary to the usual notion of the agenda as a motivator behind all the actions the agents pursue in dialogue, the private agenda here motivates only those actions the agents want to perform egoistically, with no regard for the limitations imposed upon them by the conversational setting. Cooperative agents will perform other actions that do not originate in their private agendas but that result from the obligations specified by the dialogue game. In the limit, fully cooperative agents only behave as determined by the social context. So, interviewers only ask valid questions and interviewees only respond with valid replies. These actions could also be in their private agendas, but this is accidental, as they actually result from the obligations that emerge from the dialogue game and from the analysis of these obligations given by the conversational domain module (see Sections 5.1.5 and 5.1.6). The clear case is that of interviewees having to answer a valid question by saying something that would shed a negative light upon them or their party. If they are being cooperative, they are obliged to provide the reply, but we would not expect that statement to be in their private agenda. On the other hand, the second example in Section 5.2 shows a case in which one of the messages the interviewee wants to deliver is performed as a cooperative move, just because the interviewer asked a question for which the comment worked as a valid reply. When the statement is uttered out of context, as an invalid reply motivated only by the private agenda (e.g. the first example of Section 5.2), the model does not treat this statement differently in the agenda, but it will result in a different assessment of cooperation as it constitutes a non-cooperative feature.

The function *select-goal* chooses one of the goals in the private agenda. The criteria for selecting the goal can be based on many factors. For instance, an agent could have a notion of utility associated to each goal. This would make performing goals with higher utility more desirable than those with lower utility.

In our example, agents use the following criterion for selecting private goals:

- If there are proposed questions or statements that have not yet been accepted or rejected, any accept or reject dialogue acts in the private agenda that refer to such proposals are selected (when more than one goal meets this criterion, one is selected randomly).

- Otherwise, a goal in the agenda that is not an accept or a reject dialogue act is selected at random.

Thus, for either agent $s \in \{\texttt{ir}, \texttt{ie}\}$ and an information state $\textsc{is}$, we have:

$$\textit{select-goal}_s(\textsc{is}) = \begin{cases} F(\textit{move}) \in \textsc{is}.\textsc{PrivateAgenda}, \text{if } (o, \textit{move}) \in \textsc{is}.\textsc{Proposals} \\ F(c) \in \textsc{is}.\textsc{PrivateAgenda} \text{ with } F \notin \{\textit{Accept}, \textit{Reject}\}, \text{otherwise} \\ \textit{None}, \text{ if } \textsc{is}.\textsc{PrivateAgenda} = \{\} \end{cases}$$

The function *update-agenda* reflects the effects of a dialogue move in the agent's agenda of private goals. We will use the following criterion in the example: independently of the speaker, if a dialogue act in the agent's private agenda has been performed by the agent, it is removed from the set.

So, for $s \in \{\texttt{ir}, \texttt{ie}\}$ and an information state IS, it is:

$$\textit{update-agenda}_s(\textsc{is}, (s, \textit{act})) = \textsc{is}.\textsc{PrivateAgenda} \setminus \{\textit{act}\}$$

5.1.5 Conversational Domain

The **conversational domain** allows agents to link dialogue moves with the action labels introduced in Chapter 3. This mapping can be thought of as an interpretation of the agents' contributions in the context of the current conversation and of the dialogue type[4]. The module relies on parts of the information state (e.g. the current question) and on the dialogue act performed by a speaker to give an action label. Conversely, it also encodes what dialogue acts would allow an agent to meet a given obligation in the current state of the dialogue, characterised by an information state. The module provides two functions for these purposes:

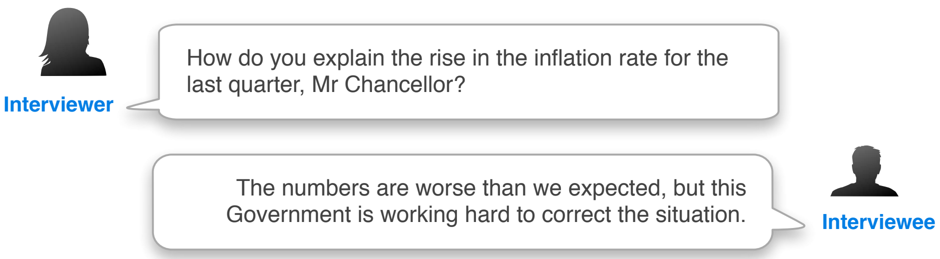
$$\begin{aligned} \textit{analyse-move} \quad &: \quad \textit{InfoState} \times Move \rightarrow \textit{Label} \\ \textit{analyse-obligation} \quad &: \quad \textit{InfoState} \times \textit{Obligation} \rightarrow \textbf{set } \textit{Act} \end{aligned}$$

The function *analyse-move* interprets a move in terms of action labels. The function *analyse-obligation* works in the opposite direction: given the current dialogue state and an obligation, i.e. a pair with a speaker and an action label, it determines what dialogue act(s) correspond to the label – and would therefore discharge the obligation when performed by the speaker. The internal details of these function depend closely on the conversational setting and on topical aspects of a particular dialogue. In a political interview setting, for instance, *analyse-move* determines whether a question is valid or invalid, whether a statement is suitable as a reply to the current

[4]This is the agents equivalent to the decisions made by annotators of the corpus study in Chapter 4, combined with the automatic mapping of the resulting annotations into action labels presented in Section 4.4.

question or just an irrelevant comment, etc. Conversely, *analyse-obligation* indicates the statement acts that would constitute a valid reply to the current question, the question acts that are valid questions, the dialogue acts that would be adequate as acceptances, etc.

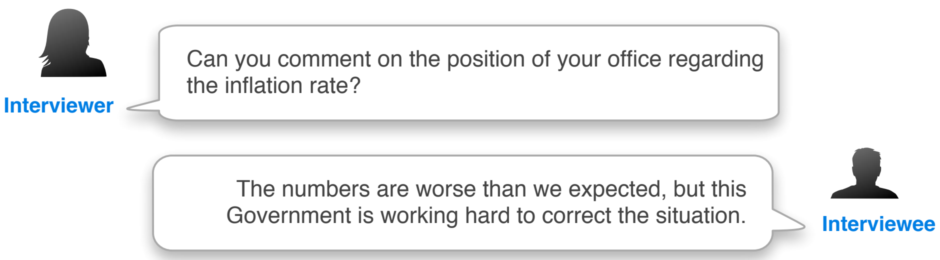
For the example model, we represent utterances as unique identifiers. Given the lack of a semantic representation, we specify *analyse-move* so that it links the identifiers contained in dialogue acts with an action label. The information state is part of the input to the function because some of the results depend on contextual aspects, such as the question currently under discussion, or the dialogue history. To illustrate this, consider the following fragment:



This pair is represented as the moves $(\texttt{ir}, \mathit{Question}(q_1))$ and $(\texttt{ie}, \mathit{Statement}(s_1))$. The interviewee's contribution is an invalid reply at this point in the dialogue (cf. Figure 3.7 in Chapter 3), so it should be:

$$\begin{aligned} \mathit{analyse\text{-}move}(\textsc{is}, (\texttt{ir}, \mathit{Question}(q_1))) &= (1) : \texttt{valid-question}\ \mathbf{N} \\ \mathit{analyse\text{-}move}(\textsc{is}, (\texttt{ie}, \mathit{Statement}(s_1))) &= (2) : \texttt{invalid-reply@}(1)\ \mathbf{I} \end{aligned}$$

Now, consider the following alternative:



Here, the moves are $(\texttt{ir}, \mathit{Question}(q_3))$ and $(\texttt{ie}, \mathit{Statement}(s_1))$. Contrary to the first example, the statement from the interviewee is a valid reply to the question and it should be:

$$\begin{aligned} \mathit{analyse\text{-}move}(\textsc{is}, (\texttt{ir}, \mathit{Question}(q_3))) &= (1) : \texttt{valid-question}\ \mathbf{N} \\ \mathit{analyse\text{-}move}(\textsc{is}, (\texttt{ie}, \mathit{Statement}(s_1))) &= (2) : \texttt{valid-reply@}(1)\ \mathbf{C} \end{aligned}$$

We achieve this is by making *analyse-move* look into the information state when classifying the move. In the first case, $\mathit{Question}(q_1)$ is the current question, for which $\mathit{Statement}(s_1)$ does not constitute a valid reply. In the second case, the current question is $\mathit{Question}(q_3)$ for which $\mathit{Statement}(s_1)$ serves as valid reply. Similarly, valid questions are determined not only by their topical relevance and adequacy in terms of the ability of the interviewee to answer them, but also by whether they have been asked before, which would be determined by the information state.

The specification of *analyse-move* for the example is as follows[5]:

$$\begin{aligned} \mathit{analyse\text{-}move}(\textsc{is}, (\texttt{ir}, \mathit{Question}(q_1))) &= \begin{cases} (i) : \texttt{valid-question}\ \mathbf{N}, \\ \quad \text{if } (\texttt{ir}, \mathit{Question}(q_1)) \notin \textsc{is}.\textsc{DialogueHistory} \\ (i) : \texttt{valid-question}\ \mathbf{R}, \text{otherwise} \end{cases} \\ \mathit{analyse\text{-}move}(\textsc{is}, (\texttt{ir}, \mathit{Question}(q_2))) &= \begin{cases} (i) : \texttt{invalid-question}\ \mathbf{N}, \\ \quad \text{if } (\texttt{ir}, \mathit{Question}(q_1)) \notin \textsc{is}.\textsc{DialogueHistory} \\ (i) : \texttt{invalid-question}\ \mathbf{R}, \text{otherwise} \end{cases} \end{aligned}$$

[5]We use (i) to denote the unique identifier generated for each action label and (j) to denote the identifier of the obvious referent label in responsive actions.

$$
\begin{aligned}
\mathit{analyse\text{-}move}(\text{IS}, (\texttt{ir}, \mathit{Question}(q_3))) &= \begin{cases} (i) : \texttt{valid-question}\ \mathbf{N}, \\ \quad \text{if } (\texttt{ir}, \mathit{Question}(q_3)) \notin \text{IS.}\textsc{DialogueHistory} \\ (i) : \texttt{valid-question}\ \mathbf{R}, \text{otherwise} \end{cases} \\
\mathit{analyse\text{-}move}(\text{IS}, (\texttt{ie}, \mathit{Statement}(s_1))) &= \begin{cases} (i) : \texttt{invalid-reply@}(j)\ \mathbf{I}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q1) \\ (i) : \texttt{invalid-reply@}(j)\ \mathbf{I}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q2) \\ (i) : \texttt{valid-reply@}(j)\ \mathbf{C}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q3) \end{cases} \\
\mathit{analyse\text{-}move}(\text{IS}, (\texttt{ie}, \mathit{Statement}(s_2))) &= \begin{cases} (i) : \texttt{valid-reply@}(j)\ \mathbf{C}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q1) \\ (i) : \texttt{invalid-reply@}(j)\ \mathbf{I}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q2) \\ (i) : \texttt{invalid-reply@}(j)\ \mathbf{I}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q3) \end{cases} \\
\mathit{analyse\text{-}move}(\text{IS}, (\texttt{ie}, \mathit{Statement}(s_3))) &= \begin{cases} (i) : \texttt{invalid-reply@}(j)\ \mathbf{I}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q1) \\ (i) : \texttt{invalid-reply@}(j)\ \mathbf{I}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q2) \\ (i) : \texttt{valid-reply@}(j)\ \mathbf{C}, \\ \quad \text{if IS.}\textsc{CurrentQuestion} = \mathit{Question}(q3) \end{cases}
\end{aligned}
$$

In addition, following the restrictions we assumed for the labels in correctly annotated political interviews, the following holds for any information state IS, speaker s, dialogue move m and content q:

$$
\begin{aligned}
\mathit{analyse\text{-}move}(\text{IS}, (s, (\mathit{Accept}(m))) &= (i) : \texttt{acceptance@}(j) \\
\mathit{analyse\text{-}move}(\text{IS}, (s, (\mathit{Reject}(m))) &= (i) : \texttt{rejection@}(j) \\
\mathit{analyse\text{-}move}(\text{IS}, (\texttt{ie}, \mathit{Question}(q)))) &= (i) : \texttt{invalid-question}
\end{aligned}
$$

Finally, as there are no statements by the interviewer that would work as prefaces to questions, the following holds for any content p:

$$\textit{analyse-move}(\textsc{is}, (\texttt{ir}, \textit{Statement}(p))) \quad = \quad (i) : \texttt{invalid-statement}$$

The function *analyse-obligation* works in the opposite direction, connecting obligations, this is pairs (s, l) where s is a speaker and l an action label, to the set of dialogue acts that meet the obligation in a given information state. For instance at the beginning of the dialogue in Figure 5.1 we have:

$$\textit{analyse-obligation}(\textsc{is}, (\texttt{ir}, \texttt{valid-question}\ \mathbf{N})) = \{\textit{Question}(q_1), \textit{Question}(q_3)\}$$

while later in the dialogue, once $\textit{Question}(q_1)$ has been asked, it is:

$$\textit{analyse-obligation}(\textsc{is}, (\texttt{ir}, \texttt{valid-question}\ \mathbf{N})) = \{\textit{Question}(q_3)\}$$

Thus, regarding new valid questions, for any information state $\textsc{is}$, it is:

$$\begin{aligned}&\textit{analyse-obligation}(\textsc{is}, (\texttt{ir}, \texttt{valid-question}\ \mathbf{N})) = \\ &\qquad \{q \in \{\textit{Question}(q_1), \textit{Question}(q_3)\} \mid (\texttt{ir}, q) \notin \textsc{is}.\textsc{DialogueHistory}\}\end{aligned}$$

As for valid complete replies, in the second turn the current questions is $\textit{Question}(q_1)$, so it should be:

$$\textit{analyse-obligation}(\textsc{is}, (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})) = \{\textit{Statement}(s_2)\}$$

In the alternative fragment, the current question is $\textit{Question}(q_3)$, so it should be:

$$\textit{analyse-obligation}(\textsc{is}, (\texttt{ie}, \texttt{valid-reply@}(1))\ \mathbf{C}) = \{\textit{Statement}(s_1), \textit{Statement}(s_3)\}$$

Thus, for any information state $\textsc{is}$, we have:

$$\begin{aligned}&\textit{analyse-obligation}(\textsc{is}, (\texttt{ie}, \texttt{valid-reply@}(j))\ \mathbf{C}) = \\ &\qquad \begin{cases}\{\textit{Statement}(s_2)\}, & \text{if } \textsc{is}.\textsc{CurrentQuestion} = \textit{Question}(q_1) \\ \{\textit{Statement}(s_1), \textit{Statement}(s_3)\}, & \text{if } \textsc{is}.\textsc{CurrentQuestion} = \textit{Question}(q_3) \\ \{\}, & \text{otherwise}\end{cases}\end{aligned}$$

The assumptions that all questions asked by interviewees are invalid and that there are no suitable statements for the interviewer in the set of available utterances mean that, for any information state IS, the following holds:

$$\begin{aligned} \textit{analyse-obligation}(\textsc{is}, (\texttt{ie}, \texttt{valid-question})) &= \{\} \\ \textit{analyse-obligation}(\textsc{is}, (\texttt{ir}, \texttt{valid-statement})) &= \{\} \end{aligned}$$

Analysing rejection and acceptance obligations, requires connecting them with the dialogue move to be accepted or rejected. The function uses IS.PROPOSALS to make this connection. The field provides a set of pairs (o, m) where o is an acceptance or rejection obligation and m a proposed question or statement move that has not yet been accepted or rejected. Thus, when analysing one of these obligations, the associated move is returned. If there are several pairs with this obligation on the same speaker, all of them are returned. Recall that the dialogue game described in Chapter 3 allows for implicit acceptances and rejections. Here we will limit ourselves to dealing with implicit rejections by repeating the current question. Implicit acceptances by asking new questions, providing replies or making statements are left for future consideration. So, for any speaker s and information state IS, it is:

$$\begin{aligned} &\textit{analyse-obligation}(\textsc{is}, (s, \texttt{acceptance@}(j))) = \\ &\qquad \{\textit{Accept}(m) \mid ((s, \texttt{acceptance@}(j))), (s, m)) \in \textsc{is}.\textsc{Proposals}\} \\ &\textit{analyse-obligation}(\textsc{is}, (s, \texttt{rejection@}(j))) = \\ &\qquad \{\textit{Reject}(m) \mid ((s, \texttt{rejection@}(j)), (s, m)) \in \textsc{is}.\textsc{Proposals}\} \quad \cup \\ &\qquad \{\textsc{is}.\textsc{CurrentQuestion} \mid \textsc{is}.\textsc{CurrentQuestion} \neq \textit{None}\} \end{aligned}$$

5.1.6 Dialogue Game

As discussed in Chapters 3 and 4, the **dialogue game** specifies the rules of the type of conversation the agents engage in. Formally, this is represented as described at the end of Chapter 3.

This module also deals with elements in the information state that are equal for both agents but specific to the conversational setting. For this purpose, it provides a function for updating the information state with the effects of a move according to the rules of the dialogue game:

$$\textit{play-move} : \textit{InfoState} \times Move \rightarrow \textit{InfoState}$$

For updating the obligations after a move, the module uses an internal function:

$$\textit{update-obligations} : \mathbf{seq}\ \textit{Obligation} \times Move \rightarrow \mathbf{seq}\ \textit{Obligation}$$

that, given a sequence of obligations and a move, follows a procedure similar to that in Algorithm 1 in Chapter 4 to introduce and discharge dynamic obligations. It uses the action label resulting from calling *analyse-move* in the conversational setting module with the move as argument.

In political interviews, *play-move* updates the obligations, the sequence of proposed questions and statements that have not yet been accepted or rejected, and the question currently under discussion.

For updating proposals after a move, the module uses an internal function:

$$\textit{update-proposals} : \mathbf{set}\ \textit{Proposal} \times Move \rightarrow \mathbf{set}\ \textit{Proposal}$$

that, given a set of proposals and a move, gives an updated set of proposals according to the dialogue game.

Asking a new question or making a statement adds a proposal to the set. This is a pair with the move and a corresponding obligation for acceptance or rejection as indicated by the dialogue game. Asking a repeated question, however, acts as an implicit rejection and removes the corresponding proposal form the set[6]. So, according to Rules (4), (7), (10), (11) and (15) of the dialogue game for political interviews (cf. Figure 3.9 in Chapter 3), for any content c, speaker s and set of proposals *props* we have[7]:

$$
\mathit{update\text{-}proposals}(\mathit{props}, (s, \mathit{Question}(c))) =
\begin{cases}
\mathit{props} \cup \{((s, \texttt{acceptance@(i)}), (s, \mathit{Question}(c))\}, \\
\qquad \text{if } \mathit{analyse\text{-}move}((s, \mathit{Question}(c))) = (i) : \texttt{valid-question}\ \mathbf{N} \\
\mathit{props} \cup \{((s, \texttt{rejection@(i)}), (s, \mathit{Question}(c))\}, \\
\qquad \text{if } \mathit{analyse\text{-}move}((s, \mathit{Question}(c))) = (i) : \texttt{invalid-question}\ \mathbf{N} \\
\mathit{props} \setminus \{((s, \texttt{rejection@(i)}), (s, \mathit{Question}(c))\}, \\
\qquad \text{if } \mathit{analyse\text{-}move}((s, \mathit{Question}(c))) = (i) : *\texttt{-question}\ \mathbf{R}
\end{cases}
$$

$$
\mathit{update\text{-}proposals}(\mathit{props}, (s, \mathit{Statement}(c))) =
\begin{cases}
\mathit{props} \cup \{((s, \texttt{acceptance@}(i)), (s, \mathit{Statement}(c))\}, \\
\qquad \text{if } \mathit{analyse\text{-}move}((s, \mathit{Statement}(c))) = (i) : \texttt{valid-reply@}(j)\ \mathbf{C} \\
\mathit{props} \cup \{((s, \texttt{rejection@}(i)), (s, \mathit{Statement}(c))\}, \\
\qquad \text{if } \mathit{analyse\text{-}move}((s, \mathit{Statement}(c))) = (i) : \texttt{invalid-reply@}(j)
\end{cases}
$$

When a move is accepted or rejected, the corresponding pair is removed from the set of pending proposals. So, for any content c, speaker s and set

[6] In future revisions, this should also consider implicit acceptances: a new question implicitly accepts a statement and a statement implicitly accepts a pending question, removing them from the list of pending proposals.

[7] As in the definition of the dialogue game in Chapter 3, the symbol `*` in an action label works as a wildcard, in this case matching valid and invalid questions.

of proposals *props*, it is:

$$\begin{aligned}
&\mathit{update\text{-}proposals}(\mathit{props}, (s, \mathit{Accept}(\mathit{move}))) = \mathit{props} \setminus \{(o, m) \in \mathit{props} | m = \mathit{move}\} \\
&\mathit{update\text{-}proposals}(\mathit{props}, (s, \mathit{Reject}(\mathit{move}))) = \mathit{props} \setminus \{(o, m) \in \mathit{props} | m = \mathit{move}\}
\end{aligned}$$

For updating the current question, the module uses a function:

$$\mathit{update\text{-}current\text{-}question} \ : \ \ \mathit{Act} \times Move \rightarrow \mathit{Act}$$

that, given the current question and a move, gives an updated current question. This field is only updated after questions and statements are accepted. When a question is accepted, it becomes the current question. So, for any content c, question q, and speakers s_1 and s_2, it is:

$$\mathit{update\text{-}current\text{-}question}(q, (s_1, \mathit{Accept}((s_2, \mathit{Question}(c))))) = \mathit{Question}(c)$$

On the other hand, when a statement is accepted as a reply to a question, the current questions is set to $None$:

$$\mathit{update\text{-}current\text{-}question}(q, (s_1, \mathit{Accept}((s_2, \mathit{Statement}(c))))) = None$$

Asking a question, making a statement or rejecting a move do not change the current question. So, for any content c, question q, speaker s and move m:

$$\begin{aligned}
\mathit{update\text{-}current\text{-}question}(q, (s, \mathit{Question}(c))) &= q \\
\mathit{update\text{-}current\text{-}question}(q, (s, \mathit{Statement}(c))) &= q \\
\mathit{update\text{-}current\text{-}question}(q, (s, \mathit{Reject}(m))) &= q
\end{aligned}$$

Using the internal functions above, *play-move* is then defined as follows:

$$
\textit{play-move}(\textsc{is}, \textit{move}) =
$$
$$
\begin{bmatrix}
\textsc{DialogueHistory} & = \textsc{is}.\textsc{DialogueHistory} \\
\textsc{PrivateAgenda} & = \textsc{is}.\textsc{PrivateAgenda} \\
\textsc{Obligations} & = \textit{update-obligations}(\textsc{is}.\textsc{Obligations}, \textit{move}) \\
\textsc{Proposals} & = \textit{update-proposals}(\textsc{is}.\textsc{Proposals}, \textit{move}) \\
\textsc{CurrentQuestion} & = \textit{update-current-question}(\textsc{is}.\textsc{CurrentQuestion}, \textit{move})
\end{bmatrix}
$$

Agents need to select a single obligation among their obligations in the sequence when deciding on their next actions. For this purpose, the module has a function:

$$
\textit{select-obligation} : \textbf{seq}\ \textit{Obligations} \times \textit{Speaker} \rightarrow \textit{Obligation}
$$

The internal details of this function depend on the particular conversational setting, as the game could define a particular order in which obligations are discharged.

In our political interviews, for instance, acceptances and rejections should be addressed before any other obligations. So, we follow this criterion:

- If there are any obligations on the speaker for acceptances or rejections, one of these is selected randomly.

- If there are any obligations on the speaker for other acts, one of these is selected randomly.

- If there are no obligations on the speakers, then $None$ is returned.

Hence, for an agent $s \in \{\mathtt{ir}, \mathtt{ie}\}$ and obligation set os, it is:

$$select\text{-}obligation(os, s) = \begin{cases} o \in \{(speaker, label) \in os \mid speaker = s \\ \quad \wedge\, label \in \{\mathtt{acceptance}, \mathtt{rejection}\}\}, \\ \quad \text{if this set is not empty} \\ o \in \{(speaker, label) \in os \mid speaker = s\} \\ \quad \text{if the above set is empty and this set is not} \\ None, \ \text{otherwise} \end{cases}$$

The module also provides a function for initialising dynamic obligations at the start of the conversation:

$$initialise\text{-}game : \mathbf{seq}\ Obligations$$

In the political interview example, there is an initial obligation on the interviewer to ask a valid question, so the following holds:

$$initialise\text{-}game(\,) \quad = \quad \langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle$$

5.1.7 Deliberation Mechanism

The **deliberation mechanism** is used by the agents to decide on what dialogue acts to perform each time they hold the floor. The module provides a function that selects the agent's next dialogue act by choosing between a private goal and an obligation:

$$select\text{-}next\text{-}act \quad : \quad Act \times Act \rightarrow Act$$

It is at this point that we can make our agents behave in different ways with respect to conversational cooperation as defined in Chapter 3. A cooperative agent would always favour obligations in the deliberation process, thus following the rules of the dialogue game. On the other hand, an agent favouring the goal dialogue acts in the private agenda will exhibit non-cooperative conversational behaviour whenever those acts are in conflict with the obligations. If the goal, interpreted under the conversational domain, does not meet an obligation, then performing that act is a non-cooperative move. Fully non-cooperative agents will always follow private goals, regardless of their obligations.

Our aim is to model agents that can exhibit behaviour ranging over different levels of cooperation. A simple way to achieve this is by allowing the deliberation mechanism to ignore obligations with probability $p \in [0, 1]$[8] each time it selects an act. The probability p is thus a parameter we can adjust to have agents behave in different ways. The function *select-next-act* will then select its first argument (the private goal) with probability p and its second argument (the obligation) with probability $(1-p)$. We achieve this by using a p-biased random binary function (a Bernoulli test with probability of success p), that will independently select either one of the arguments, each time it is called.

5.1.8 Control Algorithm

The **control algorithm** defines the interaction cycle that agents follow in a conversation. It uses the functions in the modules described above to interpret and generate utterances, update the information state, select contributions, etc. The algorithm follows essentially the same cycle in both agents:

1. Wait for an utterance from the other party and interpret it as a dia-

[8]The notation $[a, b]$ denotes the interval of real numbers between (and including) a and b, that is, the set $\{x \in \mathbb{R} \mid a \leq x \leq b\}$. Similarly, (a, b) denote the interval of real numbers between a and b but excluding these, i.e. the set $\{x \in \mathbb{R} \mid a < x < b\}$.

logue move.

2. Update the information state (both public and private).

3. Select the next dialogue move or release the floor to the other agent.

4. Update the information state (both public and private).

5. Generate the utterance corresponding to that move and send it to the other party.

In most conversational scenarios, however, there will be differences in the starting point of the cycle, as it is assumed that one of the two parties will initially hold the floor. In political interviews, for example, it is customary that the interviewer initiates the exchange. Figure 5.4 shows the control algorithms as a flow chart. For each agent, the algorithm starts with the oval labelled with the agent's name. Rectangles represent procedures, parallelograms are input/output operations and diamonds (or rhombuses) are decision points.

In order to make the interaction more realistic, we allow for several acts to be performed in each turn. Thus, when selecting the next dialogue move, agents can decide to release the floor instead of performing another act. These decisions will depend on the particular conversational setting and on the role of the agent in the dialogue. In the case of political interviews, for instance, interviewers tend to make contributions that are shorter than those of interviewees. They can also be thought of as hesitations or involuntary turn-yielding cues that allow the other agent to take the floor. In our model, we achieve this by using another biased random binary function, $release\text{-}floor : Bool$, with a parameter q that is determined empirically for

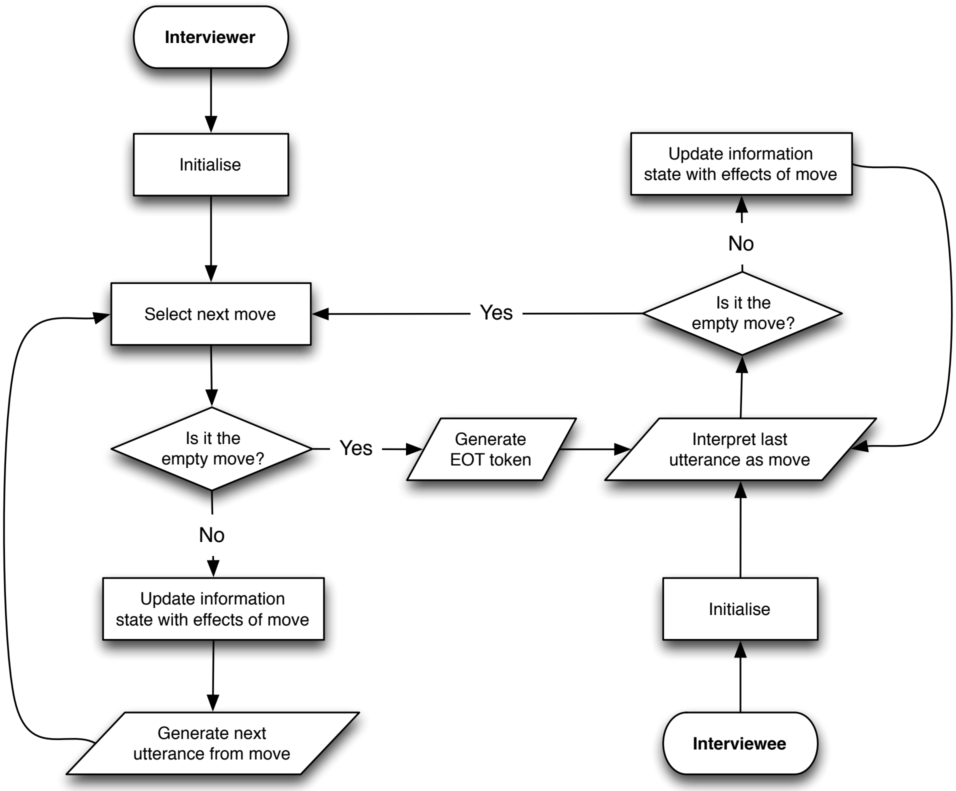


Figure 5.4: Control algorithm for political interview agents.

each agent[9]. If the trial fails, the agent chooses another act repeating steps 3-5 above. Otherwise, a special token is sent to mark the end of the turn[10]. It is worth noting that, releasing or keeping the floor is determined, either because the biased random function $release\text{-}floor$ succeeded, or because the deliberation mechanism function *select-next-act* returned $None$. The latter can happen when the deliberation favours obligations but there are none left

[9]For this function, we might prefer that the probability of success started at 0 for the first move of a turn, and that it increased after each contribution, making it less likely that the speaker kept the floor later in the turn. We can think of the parameter $q \in [0, 5]$ as the decay rate for an exponential decay function $q(k) = e^{-qk}$, using $r(k) = 1 - q(k)$ as the probability of success for a trial deciding whether to release the floor in the k-th move of a turn (taking $k = 0$ for the first move). For a discussion on this see Appendix D.

[10]The *end-of-turn token*, EOT, is a special utterance interpreted as $interpret(EOT) = (s, None)$ and generated as $generate((s, None)) = EOT$. The move $(s, None)$ is called the *empty move*.

to be discharged, or when a private goal is chosen but the private agenda is empty. A particular case for this is, for instance, when an agent is being fully non-cooperative and there are no obligations to discharge, or when it is being fully non-cooperative and it has achieved all the goals in the private agenda.

The control algorithms for the interviewer and the interviewee are shown in Algorithms 4 and 5, respectively. They include procedures corresponding roughly to steps 1-5 above, plus an initialisation. Although the control is specified as an infinite loop, we assume that a dialogue ends when both agents have produced consecutive empty turns (i.e. turns in which the only move is the empty move; see below).

The procedure INITIALISE, receives no arguments and returns a "fresh" information state with initial values for the set of obligations and the private agenda as specified, respectively, in the dialogue game and role modules of each agent:

procedure INITIALISE( )
  IS ← *initialise-state*( )
  IS.PRIVATEAGENDA ← *initialise-agenda*( )
  IS.OBLIGATIONS ← *initialise-game*( )
  return IS
end procedure

The procedure SELECT-NEXT-MOVE takes as arguments an information state and a speaker and returns the next move for the speaker. In the first place, it decides whether to release the floor using the biased random function *release-floor* introduced above. If this function fails, the procedure selects the agent's next dialogue act. To achieve this, it takes a private goal as specified by the role module. It then takes an obligation as specified by

Algorithm 4 Control algorithm for the interviewer agent.

```
IS ← INITIALISE( )
loop                                                        ▷ interaction loop starts
    next-move ← SELECT-NEXT-MOVE(IS, ir)                    ▷ speaker loop starts
    while next-move ≠ (ir, None) do
        IS ← UPDATE-INFORMATION-STATE(IS, next-move)
        next-utterance ← generate(next-move)
        next-move ← SELECT-NEXT-MOVE(IS, ir)
    end while                                               ▷ speaker loop ends
    next-utterance ← generate(next-move)                    ▷ release the floor
    last-move ← interpret(last-utterance)                   ▷ listener loop starts
    while last-move ≠ (ie, None) do
        IS ← UPDATE-INFORMATION-STATE(IS, next-move)
        last-move ← interpret(last-utterance)
    end while                                               ▷ listener loop ends
end loop                                                    ▷ interaction loop ends
```

Algorithm 5 Control algorithm for the interviewee agent.

```
IS ← INITIALISE( )
loop                                                        ▷ interaction loop starts
    last-move ← interpret(last-utterance)                   ▷ listener loop starts
    while last-move ≠ (ir, None) do
        IS ← update-information-state(IS, next-move)
        last-move ← interpret(last-utterance)
    end while                                               ▷ listener loop ends
    next-move ← select-next-move(IS, ie)                    ▷ speaker loop starts
    while next-move ≠ (ie, None) do
        IS ← update-information-state(IS, next-move)
        next-utterance ← generate(next-move)
        next-move ← select-next-move(IS, ie)
    end while                                               ▷ speaker loop ends
    next-utterance ← generate(next-move)                    ▷ release the floor
end loop                                                    ▷ interaction loop ends
```

the dialogue game and analyses it using the conversational domain module. The result is the set of dialogue acts that would discharge the selected obligation, from which one is chosen randomly[11]. Finally, it uses the deliberation mechanism to decide whether to follow the goal or to discharge the obligation. The act resulting from this decision is turned into a move and the procedure returns:

procedure SELECT-NEXT-MOVE(IS, speaker)
  if $release\text{-}floor(\,)$ **then**
    $\textit{next-move} \leftarrow (\texttt{ir}, None)$
  else
    $\textit{goal-act} \leftarrow \textit{select-goal}(\textsc{is})$
    $\textit{obligation} \leftarrow \textit{select-obligation}(\textsc{is}.\textsc{Obligations}, \textit{speaker})$
    $\textit{oligation-act} \leftarrow \textit{choice}(\textit{analyse-obligation}(\textsc{is}, \textit{obligation}))$
    $\textit{next-act} \leftarrow \textit{select-next-act}(\textit{goal-act}, \textit{obligation-act})$
    $\textit{next-move} \leftarrow (\textit{speaker}, \textit{next-act})$
  end if
  return *next-move*
end procedure

The procedure UPDATE-INFORMATION-STATE takes as arguments an information state and a move, and returns an updated information state. To update the information state, it adds the move at the end of the dialogue history sequence, updates the private agenda as specified in the role module, and the obligations, the current question and the set of pending proposals as specified in the dialogue game module:

[11]The auxiliary function *choice* : **set** $T \rightarrow T$ selects a random element from a set of elements of type T. If the set is empty, it returns $None$.

procedure UPDATE-INFORMATION-STATE(IS, move)

IS.DIALOGUEHISTORY ← IS.DIALOGUEHISTORY · $\langle move \rangle$

IS.PRIVATEAGENDA ← *update-agenda*(IS, *move*)

IS ← *play-move*(IS, *move*)

return IS

end procedure

All communication between the agents takes place through two *channels* that can be thought as blocking unidirectional pipes. From the perspective of each agent, the channels are called last-utterance and next-utterance, for input and output, respectively. Thus, agents will read their interlocutor's utterances from last-utterance and write their own utterances to next-utterance. The output channel of one agent is the input of the other, and vice-versa.

As Figure 5.4 and Algorithms 4 and 5 show, the control loop is essentially the same for both participants, except that they start at different points in the interaction cycle. The interviewer agent behaves first as a speaker, until it releases the floor, when it becomes a listener. The interviewee does this in reverse order, acting first as a listener and then, after receiving the floor, as a speaker.

5.2 Dialogue Generation and Levels of Cooperation

In this section we describe in detail how the model presented above could generate the example in Figure 5.1, hereafter referred to as Dialogue D_1. We also show how it would be possible to produce variations of this dialogue with different degrees of non-cooperation, solely by changing the parameter p in the deliberation mechanisms of both agents.

5.2.1 Generating Dialogue D_1

For generating the dialogue in Figure 5.1, we assume that both agents have a value of $p = 0.33$. This means that every time the deliberation mechanism function *select-next-act* is invoked the private goal will be selected with probability 0.33.

As specified in the agents' roles and in the dialogue game modules, in Sections 5.1.4 and 5.1.6, respectively, after the initialisation, the information state for the interviewer, $\textsc{is}_{\texttt{ir}}$, is:

$$\textsc{is}_{\texttt{ir}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle\rangle \\ \textsc{PrivateAgenda} & = & \{Question(q_2)\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & None \end{bmatrix}$$

and the information state for the interviewee, $\textsc{is}_{\texttt{ie}}$, is:

$$\textsc{is}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle\rangle \\ \textsc{PrivateAgenda} & = & \{Statement(s_1), Statement(s_3)\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & None \end{bmatrix}$$

After the initialisation, the interviewer agent enters the speaker loop and deliberates on the next move. This is the first move of the turn, so the function *release-floor* fails and a next act is selected. As there are no pending proposals, the role function *select-goal* produces the act $Question(q_2)$. The dialogue game function *select-obligation* returns $(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})$. The conversational domain function *analyse-obligation* returns the dialogue acts that, if performed, would discharge this obligation: $\{Question(q_1)$, $Question(q_3)\}$[12]. One of the two acts is chosen at random, say $Question(q_1)$,

[12]These dialogue acts come from the "stock" of available moves listed in Table 5.1. The

and the deliberation mechanism function *select-next-act* is called with arguments $Question(q_2)$ and $Question(q_1)$. The biased random test with $p = 0.33$ fails, and the second argument, that is the one resulting from the obligation, is selected as next act. The interviewer's next dialogue move is then $(\texttt{ir}, Question(q_1))$.

Once the next act is selected, the interviewer agent updates its information state to reflect the effects of this move. This involves appending the move to the dialogue history sequence and updating the private agenda, the obligations, the current question and the set of pending proposals. As $Question(q_1)$ is not in the interviewer's private agenda, this field remains the same after the update. The obligations are updated according to the dialogue game as shown in Appendix E for Turn (1)[13], so (`ir`,`valid-question` **N**) is removed from the sequence and (`ie`,`acceptance`@(1)) is added to it. Also, according to the specification of *play-move*, the current question is left unchanged, and the pair $((\texttt{ie}, \texttt{acceptance@}(1)), (\texttt{ir}, Question(q_1)))$ is added to the set of pending proposals. The resulting information state is:

$$\text{IS}_{\texttt{ir}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\texttt{ir}, Question(q_1))\rangle \\ \text{PrivateAgenda} & = & \{Question(q_2)\} \\ \text{Obligations} & = & \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle \\ \text{Proposals} & = & \{((\texttt{ie}, \texttt{acceptance@}(1)), (\texttt{ir}, Question(q_1)))\} \\ \text{CurrentQuestion} & = & None \end{bmatrix}$$

Next, the interviewer generates the utterance corresponding to the move and sends it to the interviewee through the channel `next-utterance`. As specified in Section 5.1.3, the utterance generated for the move $(\texttt{ir}, Question(q_1))$ is "How do you explain the rise in the inflation rate for the last quarter, Mr

conversational domain module has the knowledge to decide which one of such moves are valid for discharging a given obligation. Although this approach might seem artificial, in the sense that it makes agents aware of what can be said by either party at all times, it should be thought of as an ability to construct and decode utterances and to assess these utterances in terms of their validity in the current conversational context.

[13] Appendix E shows the update of obligations after each move for the entire dialogue.

Chancellor?".

In a similar vein, after the initialisation, the interviewee agent enters the listener loop and waits for an utterance from the interviewer by reading from the channel last-utterance. When the utterance is available in the channel, it is interpreted as the move $(\mathtt{ir}, \mathit{Question}(q_1))$. As this is not the empty move, the information state is updated much in the same way as for the interviewer. In this case, the private agenda does not change because the move is performed by the other agent. The resulting information state is:

$$\text{IS}_{\mathtt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \textsc{Obligations} & = & \langle(\mathtt{ie}, \mathtt{acceptance}@(1))\rangle \\ \textsc{Proposals} & = & \{((\mathtt{ie}, \mathtt{acceptance}@(1)), (\mathtt{ir}, \mathit{Question}(q_1)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

The interviewee agent then waits for another utterance to arrive on the channel.

After successfully sending the previous utterance, the interviewer selects another move. It starts by deciding whether to release the floor. As we said above, interviewer turns are usually shorter, so we use a parameter $q = 2$ as decay rate for the likelihood that the agent keeps the turn. The biased random function $\mathit{release\text{-}floor}$ then succeeds with probability 0.86 (see Appendix D for details), and the agent ends the turn. This is the case at this point in the dialogue, so the interviewer takes the empty move $(\mathtt{ir}, \mathit{None})$ as next move, generating and sending end-of-turn token EOT. The agent then enters the listener loop.

When the interviewee agent receives and interprets the end-of-turn token, it takes the floor, entering the speaker loop and selecting a next move. The agent keeps the floor and, as there are no accept or reject acts in the private agenda, the role function *select-goal* chooses one of the two statement acts at random, say, $Statement(s_1)$. The dialogue game function *select-obligation* returns $(\texttt{ie}, \texttt{acceptance}@(1))$ which, when analysed by the conversational domain function *analyse-obligation*, gives the set $\{Accept((\texttt{ir}, Question(q_1)))\}$, taking the contents of the act from the pair in the set of pending proposals, as specified in Section 5.1.5. Next, the deliberation mechanism function *select-next-act* is called with arguments $Statement(s_1)$ and $Accept((\texttt{ir}, Question(q_1)))$. The biased random test with $p = 0.33$ in *select-next-act* fails, and the second argument is selected as next act, meaning that the agent chooses to comply with the obligation. The interviewee's next dialogue move is then $(\texttt{ie}, Accept((\texttt{ir}, Question(q_1))))$.

The interviewee agent updates the information state with the effects of this move. The move is added at the end of the dialogue history and the obligations are updated, discharging the acceptance and introducing an obligation on the interviewee to provide a valid reply. The accept dialogue move also causes the pending question proposal to be removed from the set and makes $Question(q_1)$ the current question. The resulting information state is:

$$\text{IS}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, Question(q_1)), \\ & & (\texttt{ie}, Accept((\texttt{ir}, Question(q_1))))\rangle \\ \textsc{PrivateAgenda} & = & \{Statement(s_1), Statement(s_3)\} \\ \textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{valid-reply}@(1)\ \mathbf{C})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & Question(q_1) \end{bmatrix}$$

Next, the interviewee agent generates the utterance corresponding to the move and sends it to the interviewer. As specified in Section 5.1.3, the utterance generated for the move $(\texttt{ie}, Accept((\texttt{ir}, Question(q_1))))$ is "Let me try to explain this.".

The interviewer agent receives and interprets this utterance. As it is not the empty move, it updates the information state and waits for another utterance. After the update, the information state is:

$$\mathrm{IS}_{\mathtt{ir}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1)), \\ & & (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \textsc{Obligations} & = & \langle(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{Question}(q_1) \end{bmatrix}$$

The interviewee agent decides whether to release the floor. In this case, we use a parameter $q = 0.8$ for the biased random function $release\text{-}floor$, so the random function would succeed with probability 0.55 (cf. Appendix D). The trial fails, so the agent keeps the floor and selects another move. Again, the role function *select-goal* chooses the act $\mathit{Statement}(s_1)$. The dialogue game function *select-obligation* returns $(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})$ which, when analysed by the conversational domain function *analyse-obligation*, gives the set $\{\mathit{Statement}(s_2)\}$, as specified in Section 5.1.5. The deliberation mechanism function *select-next-act* is then called with arguments $\mathit{Statement}(s_1)$ and $\mathit{Statement}(s_2)$. This time, the biased random test succeeds, and the first argument, that is the private goal, is selected as the next act. The interviewee's next dialogue move is then $(\mathtt{ie}, \mathit{Statement}(s_1))$.

The interviewee agent updates the information state with the effects of this move, adding it to the dialogue history and updating the private agenda, the obligations and the set of pending proposals to include the new statement. As the statement is not a valid reply, this introduces an obligation on the interviewer to reject it (see Turn (2) in Appendix E). Also, as the statement is in the interviewee's private agenda, the act is removed from this field. The interviewee's information state results as follows:

$$
\textsc{is}_{\texttt{ie}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\
 & & \;(\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\
 & & \;(\texttt{ie}, \mathit{Statement}(s_1))\rangle \\
\textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_3)\} \\
\textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{rejection@}(3)), (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle \\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(3))), (\texttt{ie}, \mathit{Statement}(s_1)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_1)
\end{array}\right]
$$

Next, the interviewee agent generates the utterance corresponding to the move and sends it to the interviewer. As specified in Section 5.1.3, the utterance generated is "The numbers are worse than we expected, but this Government is working hard to correct the situation.". The interviewee again decides whether to release the floor. This time, the trial succeeds and the end-of-turn token is sent to the interviewer.

The interviewer receives and interprets the first utterance, updating the information state to:

$$
\textsc{is}_{\texttt{ir}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\
 & & \;(\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\
 & & \;(\texttt{ie}, \mathit{Statement}(s_1))\rangle \\
\textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\
\textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{rejection@}(3)), (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle \\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(3)), (\texttt{ie}, \mathit{Statement}(s_1)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_1)
\end{array}\right]
$$

After receiving the end-of-turn token, the interviewer agent selects a new move. Again, *select-goal* returns $\mathit{Question}(q_2)$. The function *select-obligation* returns the only obligation on the interviewer, $(\texttt{ir}, \texttt{rejection@}(3))$. As defined at the end of Section 5.1.5, the conversational domain module function *analyse-obligation* returns two dialogue acts for this obligation: an explicit rejection act and the question under discussion which works as an implicit rejection, $\{Reject((\mathtt{ie}, Statement(s_1))), Question(q_1)\}$. One of these acts is selected at random, in this case $Question(q_1)$, and passed on to the deliberation mechanism. The random trial fails in *select-next-act*, so the second act is selected, giving the move $(\mathtt{ir}, Question(q_1))$. The information state is updated with this move, discharging the obligation and removing the only element in the set of pending proposals:

$$\text{IS}_{\mathtt{ir}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\mathtt{ir}, Question(q_1)), \\ & & (\mathtt{ie}, Accept((\mathtt{ir}, Question(q_1)))), \\ & & (\mathtt{ie}, Statement(s_1)), \\ & & (\mathtt{ir}, Question(q_1))\rangle \\ \text{PrivateAgenda} & = & \{Question(q_2)\} \\ \text{Obligations} & = & \langle(\mathtt{ie}, \mathtt{valid\text{-}reply}@(1)\ \mathbf{C})\rangle \\ \text{Proposals} & = & \{\} \\ \text{CurrentQuestion} & = & Question(q_1) \end{bmatrix}$$

The interviewer generates and sends the utterance: “That’s not answering the question.”. Then, the *release-turn* random function succeeds and the interviewer releases the floor and sends the end-of-turn token.

Concurrently, the interviewee receives the first of these moves and updates the information state to:

$$\text{IS}_{\mathtt{ie}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\mathtt{ir}, Question(q_1)), \\ & & (\mathtt{ie}, Accept((\mathtt{ir}, Question(q_1)))), \\ & & (\mathtt{ie}, Statement(s_1)), \\ & & (\mathtt{ir}, Question(q_1)))\rangle \\ \text{PrivateAgenda} & = & \{Statement(s_3)\} \\ \text{Obligations} & = & \langle(\mathtt{ie}, \mathtt{valid\text{-}reply})\rangle \\ \text{Proposals} & = & \{\} \\ \text{CurrentQuestion} & = & Question(q_1) \end{bmatrix}$$

After receiving the end-of-turn token, the interviewee agent takes the floor and selects a next move. This time, *select-goal* returns $Statement(s_3)$. As before, *select-obligation* and *analyse-obligation* give $Statement(s_2)$. The deliberation mechanism function *select-next-act* is called with these two acts as arguments. The trial now fails, favouring the agent's obligation, and the second act is returned, making $(\mathtt{ie}, \mathit{Statement}(s_2))$ the next move. The agent then updates the information state, resulting in:

$$
\textsc{is}_{\mathtt{ie}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1)), \\
 & & \ (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1)))), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_1)), \\
 & & \ (\mathtt{ir}, \mathit{Question}(q_1)), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_2))\rangle \\
\textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_3)\} \\
\textsc{Obligations} & = & \langle(\mathtt{ir}, \mathtt{acceptance@}(5))\rangle \\
\textsc{Proposals} & = & \{((\mathtt{ir}, \mathtt{acceptance@}(5)), (\mathtt{ie}, \mathit{Statement}(s_2)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_1)
\end{array}\right]
$$

The utterance corresponding to this act is generated and sent to the other party: "It was due to a combination of seasonal factors and a sudden rise in the price of commodities on the international market.".

This is interpreted by the interviewer, and the information state is updated to:

$$
\textsc{is}_{\mathtt{ir}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1)), \\
 & & \ (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1)))), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_1)), \\
 & & \ (\mathtt{ir}, \mathit{Question}(q_1)), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_2))\rangle \\
\textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\
\textsc{Obligations} & = & \langle(\mathtt{ir}, \mathtt{acceptance@}(5))\rangle \\
\textsc{Proposals} & = & \{((\mathtt{ir}, \mathtt{acceptance@}(5)), (\mathtt{ie}, \mathit{Statement}(s_2)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_1)
\end{array}\right]
$$

The interviewee decides whether to release the floor. The trial fails, so the agent selects another move. The role function returns $\mathit{Statement}(s_3)$, while the dialogue game function returns $None$, as there are no obligations left on the interviewee. The deliberation mechanism is then called with $Statement(s_3)$ and $None$ as arguments. The trial succeeds, so the first one, the private goal, is selected to form the next move $(\mathtt{ie}, Statement(s_3))$. The interviewee's information state, updated with the effects of this move, is:

$$
\text{IS}_{\mathtt{ie}} = \begin{bmatrix}
\textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1)), \\
 & & \ (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1)))), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_1)), \\
 & & \ (\mathtt{ir}, \mathit{Questions}(q_1)), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_2)), \\
 & & \ (\mathtt{ie}, \mathit{Statement}(s_3))\rangle \\
\textsc{PrivateAgenda} & = & \{\} \\
\textsc{Obligations} & = & \langle(\mathtt{ir}, \mathtt{rejection@}(6)), (\mathtt{ir}, \mathtt{acceptance@}(5))\rangle \\
\textsc{Proposals} & = & \{((\mathtt{ir}, \mathtt{acceptance@}(5)), (\mathtt{ie}, \mathit{Statement}(s_2))), \\
 & & \ ((\mathtt{ir}, \mathtt{rejection@}(6)), (\mathtt{ie}, \mathit{Statement}(s_3)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_1)
\end{bmatrix}
$$

The utterance corresponding to this move is generated and sent to the interviewer agent as: "But we are working hard to correct the situation" (cf. Table 5.1). The interviewee then decides whether to release the floor. Now the trial succeeds and the end-of-turn token is sent to the interviewer.

After receiving and interpreting the first of these utterances, the interviewer updates the information state to:

$$
\textsc{is}_{\texttt{ir}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)),\\
 & & \;(\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))),\\
 & & \;(\texttt{ie}, \mathit{Statement}(s_1)),\\
 & & \;(\texttt{ir}, \mathit{Question}(q_1)),\\
 & & \;(\texttt{ie}, \mathit{Statement}(s_2)),\\
 & & \;(\texttt{ie}, \mathit{Statement}(s_3))\rangle\\
\textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\}\\
\textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ir}, \texttt{acceptance@}(5))\rangle\\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{acceptance@}(5)), (\texttt{ie}, \mathit{Statement}(s_2))),\\
 & & \;((\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ie}, \mathit{Statement}(s_3)))\}\\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_1)
\end{array}\right]
$$

The end-of-turn token follows, so the interviewer takes the floor and select a new move. The role module gives the goal $\mathit{Question}(q_2)$ and *select-obligation* chooses one of the two obligations on the interviewer, $(\texttt{ir}, \texttt{acceptance@}(5))$ for instance, which is analysed as the set $\{\mathit{Accept}(\texttt{ie}, \mathit{Statements}(s_2))\}$. The deliberation mechanism is then called with arguments $\mathit{Question}(q_2)$ and $\mathit{Accept}(\texttt{ie}, \mathit{Statements}(s_2))$. The random function fails and the second act, that is the obligation, is selected, making $(\texttt{ir}, \mathit{Accept}(\texttt{ie}, \mathit{Statements}(s_2)))$ the next move. The updated information state for the interviewer is:

$$
\textsc{is}_{\texttt{ir}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)),\\
 & & \;(\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))),\\
 & & \;(\texttt{ie}, \mathit{Statement}(s_1)),\\
 & & \;(\texttt{ir}, \mathit{Question}(q_1)),\\
 & & \;(\texttt{ie}, \mathit{Statement}(s_2)),\\
 & & \;(\texttt{ie}, \mathit{Statement}(s_3)),\\
 & & \;(\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2))))\rangle\\
\textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\}\\
\textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N}), (\texttt{ir}, \texttt{rejection@}(6))\rangle\\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ie}, \mathit{Statement}(s_3)))\}\\
\textsc{CurrentQuestion} & = & \mathit{None}
\end{array}\right]
$$

The utterance for this move is generated and sent as "That's clear enough.". The function *release-floor* fails, so the interviewer selects another move. The role gives again the goal $Question(q_2)$. The function *select-obligation* returns $(\texttt{ir}, \texttt{rejection@}(6))$, the only obligation on the interviewer, which is analysed by the conversational domain. As the current question is $None$, the analysis results in $\{Reject((\texttt{ie}, Statement(s_3))\}$. The deliberation mechanism is called and the trial succeeds favouring the private goal. The interviewer's next move is $(\texttt{ir}, Question(q_2))$ and the information state update results in:

$$
\textsc{is}_{\texttt{ir}} = \left[\begin{array}{lcl}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, Question(q_1)), \\
 & & (\texttt{ie}, Accept((\texttt{ir}, Question(q_1)))), \\
 & & (\texttt{ie}, Statement(s_1)), \\
 & & (\texttt{ir}, Question(q_1)), \\
 & & (\texttt{ie}, Statement(s_2)), \\
 & & (\texttt{ie}, Statement(s_3)), \\
 & & (\texttt{ir}, Accept((\texttt{ie}, Statement(s_2)))), \\
 & & (\texttt{ir}, Question(q_2))\rangle \\
\textsc{PrivateAgenda} & = & \{\} \\
\textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{rejection@}(8)), (\texttt{ir}, \texttt{valid-question}\ \mathbf{N}), \\
 & & (\texttt{ir}, \texttt{rejection@}(6))\rangle \\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ie}, Statement(s_3))), \\
 & & ((\texttt{ie}, \texttt{rejection@}(8)), (\texttt{ir}, Question(q_2)))\} \\
\textsc{CurrentQuestion} & = & None
\end{array}\right]
$$

The utterance is generated and sent as "Could you please tell us the price of a pint of milk in China?". The interviewer then deliberates on whether to release the floor, which succeeds, so the end-of-turn token is sent to the other agent.

Upon receiving the first move, the interviewee agent interprets it and updates the information state to:

$$\textsc{is}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_1)), \\ & & (\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ie}, \mathit{Statement}(s_3)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2))))\rangle \\ \textsc{PrivateAgenda} & = & \{\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question } \mathbf{N}), (\texttt{ir}, \texttt{rejection@}(6))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ie}, \mathit{Statement}(s_3)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

After the second move, the interviewee's information state is :

$$\textsc{is}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_1)), \\ & & (\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ie}, \mathit{Statement}(s_3)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & (\texttt{ir}, \mathit{Question}(q_2))\rangle \\ \textsc{PrivateAgenda} & = & \{\} \\ \textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{rejection@}(8)), (\texttt{ir}, \texttt{valid-question } \mathbf{N}), \\ & & (\texttt{ir}, \texttt{rejection@}(6))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ie}, \mathit{Statement}(s_3))), \\ & & ((\texttt{ie}, \texttt{rejection@}(8)), (\texttt{ir}, \mathit{Question}(q_2)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

With the end-of-turn token, the interviewee takes the floor, selecting a next move. As the private agenda is empty, *select-goal* returns $None$. The dialogue game function *select-obligation* returns $(\texttt{ie}, \texttt{rejection@}(8))$, which analysed by the conversational domain gives $\{\mathit{Reject}((\texttt{ir}, \mathit{Question}(q_2)))\}$. The deliberation mechanism is called and the rejection is selected, giving

$(\texttt{ie}, \mathit{Reject}((\texttt{ir}, \mathit{Question}(q_2))))$ as the next move:

$$\text{IS}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_1)), \\ & & (\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ie}, \mathit{Statement}(s_3)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & (\texttt{ir}, \mathit{Question}(q_2)), \\ & & (\texttt{ie}, \mathit{Reject}((\texttt{ir}, \mathit{Question}(q_2))))\rangle \\ \textsc{PrivateAgenda} & = & \{\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question }\mathbf{N}), (\texttt{ir}, \texttt{rejection@}(6))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(6)), (\texttt{ie}, \mathit{Statement}(s_3)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

The corresponding utterance is generated and sent to the interviewer as "I suppose you don't expect me to answer that question.". The agent then decides whether to release the floor, which succeeds, so the end-of-turn token is sent to the interviewer.

The interviewer receives the first utterance and updates the information state:

$$\text{IS}_{\texttt{ir}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_1)), \\ & & (\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ie}, \mathit{Statement}(s_3)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & (\texttt{ir}, \mathit{Question}(q_2)), \\ & & (\texttt{ie}, \mathit{Reject}((\texttt{ir}, \mathit{Question}(q_2))))\rangle \\ \textsc{PrivateAgenda} & = & \{\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question }\mathbf{N}), (\texttt{ir}, \texttt{rejection@}(1))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ir}, \texttt{rejection@}(1)), (\texttt{ie}, \mathit{Statement}(s_3)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

After receiving the end-of-turn token, the interviewer takes the floor and selects a next move. The function *select-goal* gives $None$. As before, *select-obligation* returns $(\texttt{ir}, \texttt{rejection}@(6))$[14], which analysed by the domain becomes the set $\{Reject((\texttt{ie}, Statement(s_3)))\}$. The deliberation mechanism again favours the goal, $None$, which becomes the end-of-turn move $(\texttt{ir}, None)$. There is no information state update and the end-of-turn token is sent to the interviewer.

Upon receiving this the interviewer selects a new move. As the agenda is empty and there are no obligations on the interviewee, the deliberation mechanism is called with two $None$ arguments. Regardless of the outcome of the random trial, the next move is the empty move and the agent releases the floor.

As the last two turns were empty turns, the dialogue ends.

5.2.2 Generating a Fully Cooperative Alternative

Let us now see how we to generate the dialogue in Figure 5.5, in which both participants follow the rules of the game. We make both agents fully cooperative we set $p = 0$ in their deliberation mechanism. This means that the function *select-next-act* will always choose the argument corresponding to an obligation when selecting a dialogue act.

[14]This option might appear unrealistic at this stage in the dialogue. The invalid reply is quite far back in the dialogue history what could affect the coherence of a rejection at this point – although good phrasing could make it fit naturally (e.g. "Your earlier remark on the position of the Government was not dealing with my question, but could you please answer the following..."). An alternative would be to incorporate a dialogue history threshold as we did in Chapter 4, letting agents disregard their obligations after these have been in IS.OBLIGATIONS for a certain number of turns.

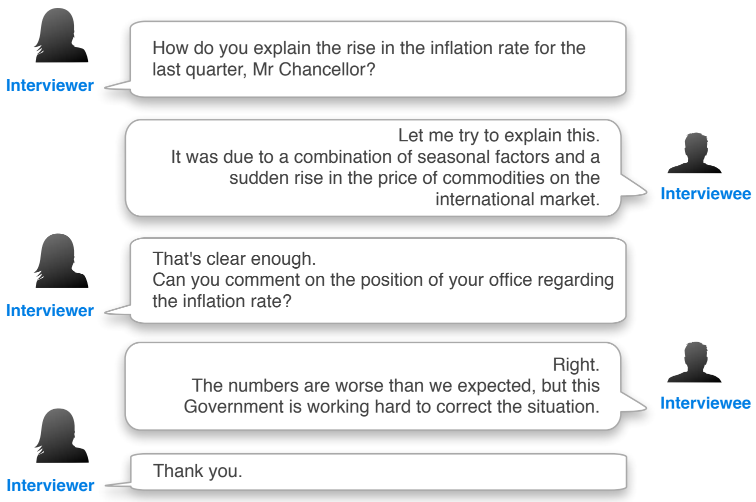


Figure 5.5: A fully cooperative political interview (D_2).

As before, after initialisation, the agents' information states are:

$$\text{IS}_{\texttt{ir}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle\rangle \\ \text{PrivateAgenda} & = & \{Question(q_2)\} \\ \text{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle \\ \text{Proposals} & = & \{\} \\ \text{CurrentQuestion} & = & None \end{bmatrix}$$

$$\text{IS}_{\texttt{ie}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle\rangle \\ \text{PrivateAgenda} & = & \{Statement(s_1), Statement(s_3)\} \\ \text{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle \\ \text{Proposals} & = & \{\} \\ \text{CurrentQuestion} & = & None \end{bmatrix}$$

The interviewer then select the next act between $Question(q_2)$ and $Question(q_1)$. The random trial with $p = 0$ fails in *select-next-act* , favouring the obligation, and the agent sends the utterance: "How do you explain the rise in the inflation rate for the last quarter, Mr Chancellor?"

and releases the floor. The updated information states are:

$$\text{IS}_{\texttt{ir}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1))\rangle \\ \text{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \text{Obligations} & = & \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle \\ \text{Proposals} & = & \{((\texttt{ie}, \texttt{acceptance@}(1)), (\texttt{ir}, \mathit{Question}(q_1)))\} \\ \text{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

$$\text{IS}_{\texttt{ie}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1))\rangle \\ \text{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \text{Obligations} & = & \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle \\ \text{Proposals} & = & \{((\texttt{ie}, \texttt{acceptance@}(1)), (\texttt{ir}, \mathit{Question}(q_1)))\} \\ \text{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

Now, the interviewee agent selects a dialogue act between $\mathit{Statement}(s_1)$ and $\mathit{Accept}(\texttt{ir}, \mathit{Question}(q_1))$. The random trial fails in *select-next-act* and the second act is chosen for the next move, favouring the agent's obligation. The agent utters "Let me try to explain this." and the updated information states are:

$$\text{IS}_{\texttt{ir}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \quad (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1))))\rangle \\ \text{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \text{Obligations} & = & \langle(\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle \\ \text{Proposals} & = & \{\} \\ \text{CurrentQuestion} & = & \mathit{Question}(q_1) \end{bmatrix}$$

$$\text{IS}_{\texttt{ie}} = \begin{bmatrix} \text{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \quad (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1))))\rangle \\ \text{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \text{Obligations} & = & \langle(\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle \\ \text{Proposals} & = & \{\} \\ \text{CurrentQuestion} & = & \mathit{Question}(q_1) \end{bmatrix}$$

The interviewee keeps the floor, selecting a new act. The deliberation mechanism is called with arguments $\mathit{Statement}(s_3)$ and $\mathit{Statement}(s_2)$. The second is selected and the agent utters "It was due to a combination of seasonal factors and a sudden rise in the price of commodities on the international market.". The updated information states are:

$$\mathrm{IS}_{\texttt{ir}} = \left[\begin{array}{lcl} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & \ (\texttt{ie}, \mathit{Statement}(s_2))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{acceptance@}(3))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ir}, \texttt{acceptance@}(3)), (\texttt{ie}, \mathit{Statement}(s_2)))\} \\ \textsc{CurrentQuestion} & = & \mathit{Question}(q_1) \end{array}\right]$$

$$\mathrm{IS}_{\texttt{ie}} = \left[\begin{array}{lcl} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & \ (\texttt{ie}, \mathit{Statement}(s_2))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{acceptance@}(3))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ir}, \texttt{acceptance@}(3)), (\texttt{ie}, \mathit{Statement}(s_2)))\} \\ \textsc{CurrentQuestion} & = & \mathit{Question}(q_1) \end{array}\right]$$

The interviewee releases the floor and the interviewer selects a next act between $\mathit{Question}(q_2)$ and $\mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2))$. The second is chosen and the interviewer utters "That's clear enough.", with the resulting information states:

$$\mathrm{IS}_{\texttt{ir}} = \left[\begin{array}{lcl} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & \ (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & \ (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{array}\right]$$

$$\mathrm{IS}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{valid-question}\mathbf{N})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

The interviewer keeps the floor and chooses the second of $\mathit{Question}(q_2)$ and $\mathit{Question}(q_3)$, uttering "Can you comment on the position of your office regarding the inflation rate?". The information states are:

$$\mathrm{IS}_{\texttt{ir}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & (\texttt{ir}, \mathit{Question}(q_3))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{acceptance@}(5))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ie}, \texttt{acceptance@}(5)), (\texttt{ir}, \mathit{Question}(q_3)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

$$\mathrm{IS}_{\texttt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & (\texttt{ir}, \mathit{Question}(q_3))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{acceptance@}(5))\rangle \\ \textsc{Proposals} & = & \{((\texttt{ie}, \texttt{acceptance@}(5)), (\texttt{ir}, \mathit{Question}(q_3)))\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

The floor is released and the interviewee selects the next contribution between $\mathit{Statement}(s_3)$ and $\mathit{Accept}((\texttt{ir}, \mathit{Question}(q_3)))$. The second is chosen and the agent utters "Right.". The resulting information states are:

$$\mathrm{IS}_{\texttt{ir}} = \left[\begin{array}{lcl} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & \ (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & \ (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & \ (\texttt{ir}, \mathit{Question}(q_3)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_3))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{valid-reply@}(6)\ \mathbf{C})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{Question}(q_3) \end{array}\right]$$

$$\mathrm{IS}_{\texttt{ie}} = \left[\begin{array}{lcl} \textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\ & & \ (\texttt{ie}, \mathit{Statement}(s_2)), \\ & & \ (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\ & & \ (\texttt{ir}, \mathit{Question}(q_3)), \\ & & \ (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_3))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_1), \mathit{Statement}(s_3)\} \\ \textsc{Obligations} & = & \langle(\texttt{ie}, \texttt{valid-reply@}(6)\ \mathbf{C}))\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{Question}(q_3) \end{array}\right]$$

The interviewer keeps the floor and chooses another act between $Statement(s_3)$ and $Statement(s_1)$. The second is selected by the deliberation mechanism and the agent utters "The numbers are worse than we expected, but this Government is working hard to correct the situation.".

The information states after this move are:

$$
\text{IS}_{\texttt{ir}} = \begin{bmatrix}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\
 & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\
 & & (\texttt{ie}, \mathit{Statement}(s_2)), \\
 & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\
 & & (\texttt{ir}, \mathit{Question}(q_3)), \\
 & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_3)))), \\
 & & (\texttt{ie}, \mathit{Statement}(s_1))\rangle \\
\textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\
\textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{acceptance@}(7))\rangle \\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{acceptance@}(7)), (\texttt{ie}, \mathit{Statement}(s_1)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_3)
\end{bmatrix}
$$

$$
\text{IS}_{\texttt{ie}} = \begin{bmatrix}
\textsc{DialogueHistory} & = & \langle(\texttt{ir}, \mathit{Question}(q_1)), \\
 & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_1)))), \\
 & & (\texttt{ie}, \mathit{Statement}(s_2)), \\
 & & (\texttt{ir}, \mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_2)))), \\
 & & (\texttt{ir}, \mathit{Question}(q_3)), \\
 & & (\texttt{ie}, \mathit{Accept}((\texttt{ir}, \mathit{Question}(q_3)))), \\
 & & (\texttt{ie}, \mathit{Statement}(s_1))\rangle \\
\textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_3)\} \\
\textsc{Obligations} & = & \langle(\texttt{ir}, \texttt{acceptance@}(7))\rangle \\
\textsc{Proposals} & = & \{((\texttt{ir}, \texttt{acceptance@}(7)), (\texttt{ie}, \mathit{Statement}(s_1)))\} \\
\textsc{CurrentQuestion} & = & \mathit{Question}(q_3)
\end{bmatrix}
$$

Note that in this case, one of the elements of the interviewee's private agenda was removed, as the goal was achieved by following the obligation.

The interviewer releases the floor and the interviewee selects a next move. The private goal is again $\mathit{Question}(q_2)$ and the obligation act $\mathit{Accept}((\texttt{ie}, \mathit{Statement}(s_1)))$. The second is chosen and the interviewer utters "Thank you." and releases the floor. The information states are:

$$\mathrm{IS}_{\mathtt{ir}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1)), \\ & & (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1)))), \\ & & (\mathtt{ie}, \mathit{Statement}(s_2)), \\ & & (\mathtt{ir}, \mathit{Accept}((\mathtt{ie}, \mathit{Statement}(s_2)))), \\ & & (\mathtt{ir}, \mathit{Question}(q_3)), \\ & & (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_3)))), \\ & & (\mathtt{ie}, \mathit{Statement}(s_1)) \\ & & (\mathtt{ir}, \mathit{Accept}((\mathtt{ie}, \mathit{Statement}(s_1))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Question}(q_2)\} \\ \textsc{Obligations} & = & \langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

$$\mathrm{IS}_{\mathtt{ie}} = \begin{bmatrix} \textsc{DialogueHistory} & = & \langle(\mathtt{ir}, \mathit{Question}(q_1)), \\ & & (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_1)))), \\ & & (\mathtt{ie}, \mathit{Statement}(s_2)), \\ & & (\mathtt{ir}, \mathit{Accept}((\mathtt{ie}, \mathit{Statement}(s_2)))), \\ & & (\mathtt{ir}, \mathit{Question}(q_3)), \\ & & (\mathtt{ie}, \mathit{Accept}((\mathtt{ir}, \mathit{Question}(q_3)))), \\ & & (\mathtt{ie}, \mathit{Statement}(s_1)) \\ & & (\mathtt{ir}, \mathit{Accept}((\mathtt{ie}, \mathit{Statement}(s_1))))\rangle \\ \textsc{PrivateAgenda} & = & \{\mathit{Statement}(s_3)\} \\ \textsc{Obligations} & = & \langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle \\ \textsc{Proposals} & = & \{\} \\ \textsc{CurrentQuestion} & = & \mathit{None} \end{bmatrix}$$

The interviewee takes the floor and chooses the next act between $\mathit{Statement}(s_3)$ and, as there are no obligations to discharge, None, selecting the second and sending the end-of-turn token to the other agent. Then, the interviewer chooses between $\mathit{Question}(q_2)$ and, as there are no valid questions left in the conversational domain, None, selecting the second and sending the end-of-turn token to the interviewee. As turns were empty twice in a row, the dialogue ends.

5.2.3 A Note on p and the Degree of Non-Cooperation

Following the process described above, it would be possible to generate dialogues in which the agents exhibit different levels of cooperation by changing the value of the parameter p. For instance, to generate an interview similar to dialogue D_3 (cf. Section 5.1.6), in which the interviewer is fully cooperative but the interviewee does not always follow the rules of the dialogue game, we would set $p = 0$ for the interviewer agent and $p = 0.5$ for the interviewee agent.

Moreover, the values for p not necessarily need to remain constant throughout the dialogue. By incorporating the mechanism for measuring cooperation that we presented in Chapter 4, at any point in the dialogue, agents could assess each other's behaviour and adjust their own p values accordingly[15]. This would allow, for example, the implementation of a *tit-for-tat* strategy, in which an agent starts behaving cooperatively ($p = 0$) and later on, e.g. after every other turn, adopts the degree of non-cooperation of the other party as the new value for p. A similar approach would allow us to simulate *escalation*, a situation in which both agents adjust the value of p to be slightly higher than the degree of non-cooperation of the other party.

With such a mechanism, agents could also measure their own degree of non-cooperation, and make adjustments accordingly. As we use random probabilistic functions, it would be possible for an agent to act significantly more or less cooperatively than the value suggested by p, especially in rather short exchanges in which the probability distribution of the actual random trials might not be sufficiently close to that indicated by the parameter.

[15] As the interpretation of moves in terms of actions levels depends on the current information state, agents would also have to keep the action labels corresponding to the moves they perform as part of the dialogue history. These would then be readily available as input to the mechanism for measuring de degree of non-cooperation.

Agents could then increase or decrease the value of p dynamically to allow for adjustments.

5.3 Comparison with Related Approaches to Modelling Conversational Agents

The modelling approach for conversational agents described and illustrated in the previous sections bears some similarities with previous research:

- The interaction structured in terms of the performance of dialogue acts follows on Traum and Hinkelman's (1992) **conversation acts**, on the dialogue acts of the **Discourse Resource Initiative** (Initiative, 1997; Poesio and Traum, 1998) and on the **ISO 24617-2** standard recently proposed by Bunt and others (Bunt et al., 2012). These are extensions of speech act theory (Searle, 1969) to account for turn-taking, grounding, the exchange of information, the management of underlying task, etc.

- The use of a central structure, the information state, to keep track of the informational elements throughout the conversation can be traced back to Poesio and Traum's **conversational score** (Poesio and Traum, 1997; Poesio and Traum, 1998), to Ginzburg's **dialogue gameboard** (Ginzburg, 1996; Ginzburg, 1997; Ginzburg, 2012), to Piwek's (1998) **conversational store**, to Beun's **cognitive state** – which he also refers to as **game-board** (Beun, 2001) –, and more closely to the **information state update** tradition put forward in the context of the TRINDI project (Poesio et al., 1999; Traum et al., 1999; Larsson and Traum, 2000; Traum and Larsson, 2003; Kreutel and Matheson, 2003b).

- The notion of a current question to condition the adequacy of replies is distantly inspired by Ginzburg's **question under discussion** (Ginzburg, 1997; Ginzburg, 2012).

- The structure and management of proposals, introduced by one of the dialogue participants and either accepted or rejected by the other is based on Clark's **joint actions**, more specifically on those at level 4 of the **action ladder**, that is the notion of **joint projects** (Clark, 1996, Chapter 7).

- The focus on discourse obligations to deal with non-cooperation in dialogue was first proposed by Traum and Allen (1994). Discourse obligations are part of **PTT**, the theory of dialogue developed by Poesio and Traum (1997; 1998) and partially implemented in the EDIS dialogue system (Matheson et al., 2000), with emphasis on grounding. Discourse obligations are central to **obligation-driven dialogue modelling** (Kreutel and Matheson, 1999; Kreutel and Matheson, 2000; Kreutel and Matheson, 2003b) where they are proposed as the primitives from which other structures such as intentions and local dialogue games can be derived.

Nonetheless, our modelling approach departs from those mentioned above in several key aspects:

- The taxonomy of dialogue acts is minimalistic in comparison to those in the Discourse Resource Initiative, Poesio and Traum's theory and Bunt and others' ISO standard. Instead of a hierarchy in which a move can be decomposed into several moves at lower levels of interaction, we consider a flat taxonomy of the dialogue acts for the purpose of analysing the dynamics of discourse obligations. This is even simpler than the taxonomy we used in Chapter 4 to annotate the corpus of political interviews and the one used in the prototypes described in the next section. The lack of an explicit semantic representation for the content of dialogue acts and the omission of responsive dialogue acts beyond acceptances and rejections are the main differences. This is in line with the simplifying assumptions listed at the beginning of Section 5.1 and compensated by the use of action labels that operate as pragmatic interpretations of the performance of these simple dialogue act in context.

- In the information state update tradition, state changes are specified by a collection of independent update rules, whose preconditions activate after certain dialogue moves and in certain values of the information state. A second set of selection rules then determines which rules are applied from the set of active rules until no rules can be applied. This approach is well-suited for the construction of experimental dialogue system, allowing for incremental escalation and improvement by the addition, subtraction or modification of these independent rules. However, the resulting dynamics is harder to describe and analyse, as the extent to which these rules interfere with each other is difficult to trace – something noted by Bos et al. (2003) when describing DIPPER, a Java reimplementation of the TrindiKit information update engine that dispenses with selection rules. We abandoned the use of independent rules to update the information state in favour of a centralised control algorithm that determined the sequence in which the functions in each module are activated.

- While Ginzburg's question under discussion (QUD) is represented as a stack, allowing for several questions to be open for discussion at the same time, our current question holds a single value. The mechanism by which questions are placed and removed from this field is also considerably simpler than Ginzburg's. In our model a question becomes the current question only after it is proposed by one of the parties and accepted by the other. The question ceases to be the current questions either by being resolved, that is after a statement by one of the parties is accepted as a reply by the other participant, or by being replaced by a new question. This is considerably simpler than Ginzburg's QUD-DOWNDATING, especially as acceptances and rejections are not negotiated in our model, in part to prevent potentially infinite chains (e.g. the acceptance of a rejection of an acceptance of an acceptance. . .).

- Clark's joint actions ladder considers four levels of interaction, three of which are not part of our model as per the assumptions listed in Section 5.1 that the coding and decoding of the participant's utterances in terms of dialogue acts is flawless (so no grounding mechanisms are needed) and that the conversational domain and the dialogue game are equal, so the interpretations of those dialogue acts in terms of action labels and discourse obligations is the same for both participants. It is only at level four, that of proposing and taking up or declining a joint project that our agents can operate. Clark (1996, Chapter 7) considers four possible outcomes to the joint actions in level 4:

  1. Success: the hearer takes up on the proposed task
  2. Partial success: the hearer takes up on an alternative project
  3. Failure: the hearer declines
  4. Breakdown: the hearer abandons the exchange

The second of these outcomes is not explicitly part of our model, but could result, for instance from hearers accepting a proposal but then failing to carry out their part of the task (e.g. an interviewer accepting a question but failing to provide a valid answer).

- Our treatment of discourse obligations is more straightforward than in Poesio and Traum's theory (Poesio and Traum, 1998) and in the implementation as a dialogue system described by Matheson et al. (2000). In their approach, obligations are passed through a grounding mechanism in the same way as dialogue acts. Instead, in our approach they are outside the scope of any grounding mechanisms: discourse obligations are a direct consequence of the dialogue acts performed by the speakers and operate immediately once those dialogue acts have been grounded. Kreutel and Matheson (2003b) also dispense with grounding obligations, although they suggest that a future extension of their approach would incorporate grounding mechanism in the tradition of Poesio and Traum's theory. A second simplification with respect to these authors is the absence of conditional updates that can introduce obligations based on the performance of subsequent dialogue acts. This effect is instead specified in the rules of the dialogue game which in turn is the most salient difference between our treatment of obligations and that in the work mentioned above. All the research following Traum and Allen's (1994) introduction of discourse obligations has regarded the rules that govern the introduction and discharge of obligations as part of the theory of dialogue under consideration. As far as we know, this is the first approach in which these rules are made explicitly dependent of the type of interaction and detached from other aspects of conversational interaction.

5.4 Prototype Implementation

In devising the concepts and mechanisms presented in this chapter and many of those discussed earlier in the thesis, we implemented a prototype of non-cooperative conversational agents that can interact with each other and also separately with a human user simulating a political interview.

The agents borrow some elements from `py-trindikit` (Ljunglöf, 2009), a Python implementation of the TrindiKit dialogue toolkit (Larsson et al., 2000; Larsson and Traum, 2000). Changes were made so that two agents could interact with each other, as well as with the user as originally intended in TrindiKit. The system consists of three modules:

nca_types.py: defines basic data structures (e.g. stack, stackset) and types for representing semantic and pragmatic data, independent of the dialogue theory (e.g. Move, Question, Prop). Follows closely the implementation of their elements in `py-trindikit` (Ljunglöf, 2009).

nca.py: defines abstract and domain-independent aspects of a theory of conversational agents able to display and deal with non cooperative behaviour[16]: InfoState, Game, Domain, Deliberation, Agent (Role) and Moves (Dialogue Act Taxonomy). The dialogue acts are organised as follows:

- Turn taking: Greet, Quit
- Core: Ask (containing a Question), Assert (containing a Prop), Answer (containing a Question), Address (Reject or Accept, containing another move)

[16] These roughly correspond to the elements described in Section 5.1 given in parenthesis when the name differs significantly.

twist.py : implements the system, with concrete classes for the game (Interview), which is shared by both agents and specifies what happens after each move, plus other domain-specific aspects (e.g. relevance of answers with respect to questions, the topical agenda); the roles (Interviewer and Interviewee), with a specification of the control sequence and instances of the deliberation mechanism; the user (User), to allow for interaction with either of the pre-defined roles; and the *main* function that creates two child processes running each participant and connects them through bidirectional pipes.

The level of cooperation of each agent is chosen by setting a runtime parameter called `lnc` that can take 5 possible values:

- lnc = 0, favouring obligations
- lnc = 1, random choice with bias towards obligations
- lnc = 2, random choice without bias
- lnc = 3, random choice with bias towards private goals
- lnc = 4, favouring private goals

The implementation of the agents departs from the information state update approach in that the changes in the state are controlled explicitly, rather than by means of update rules that are triggered automatically by the dialogue management engine once their preconditions are met. However, it is possible to put these in terms of transition rules in a finite state machine as sketched in Figures 5.6 and 5.7. The notation is as follows:

- Green states are initial.
- Red states are final.

- Blue states are deliberation points in which the agent must decide between favouring obligations or private goals.

- Preconditions to a transition appear before the symbol `--->` and are dialogue acts performed by others or auxiliary functions that evaluate to a boolean value.

- Effects of transitions appear after the symbol `--->` and can be dialogue acts performed by self, auxiliary functions or changes to the information state (obligations only).

The code of the current prototype system is available from the author. Future work includes updating it to include all the elements in the model described above and an analysis-by-synthesis study on the resulting simulations, to assess the extent to which the degree of cooperation effectively grows as we increase the value of the runtime parameters that control the level to which each agent favours discourse obligations over private goals.

5.5 Summary

This chapter proposed a model of conversational agents that can deal and exhibit linguistic behaviour with varying degrees of cooperation. With a focus on dialogue management, we achieve this by incorporating and reasoning about the concepts, structures and algorithms introduced and evaluated earlier in the thesis.

We started with a description of the elements that make up the cognitive architecture of the agents and explained the simplifying assumptions that we made in order to focus the discussion on the management of cooperation. Several lines of future works result from relaxing one or more of these

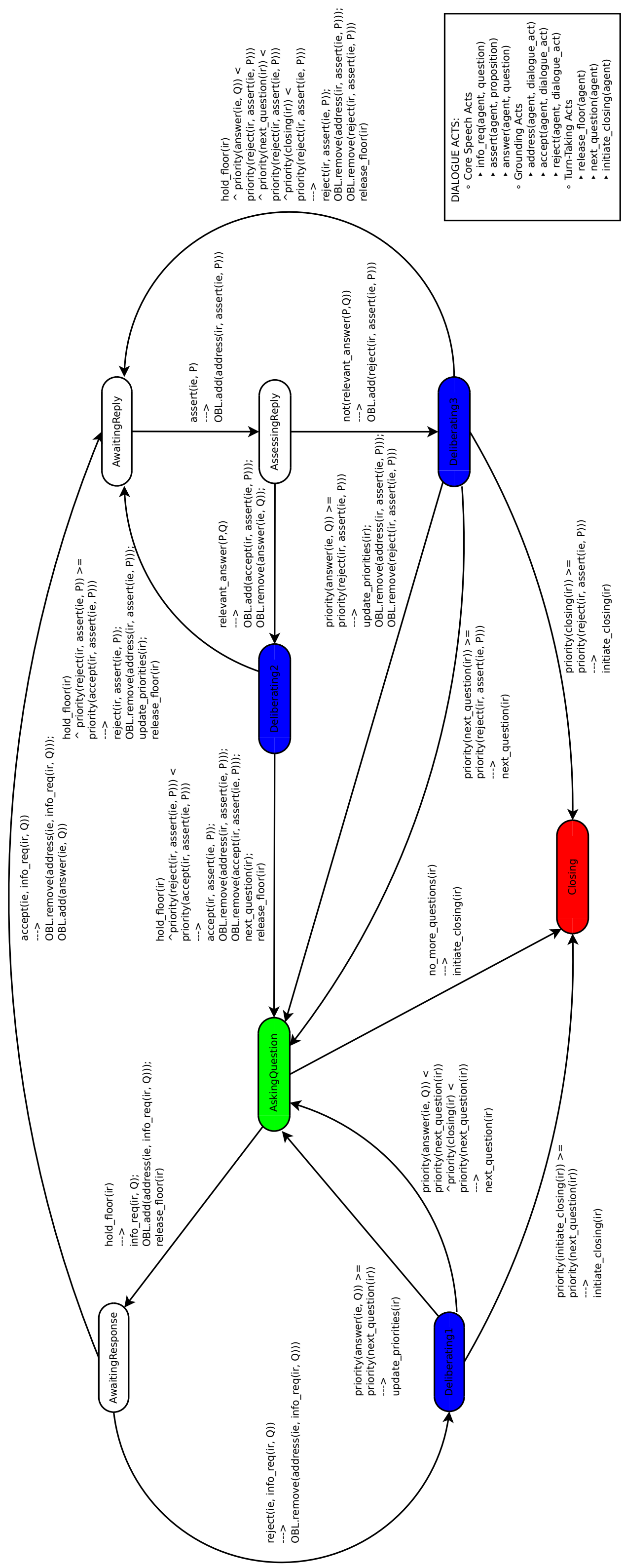


Figure 5.6: A finite state machine specifying the interviewer prototype

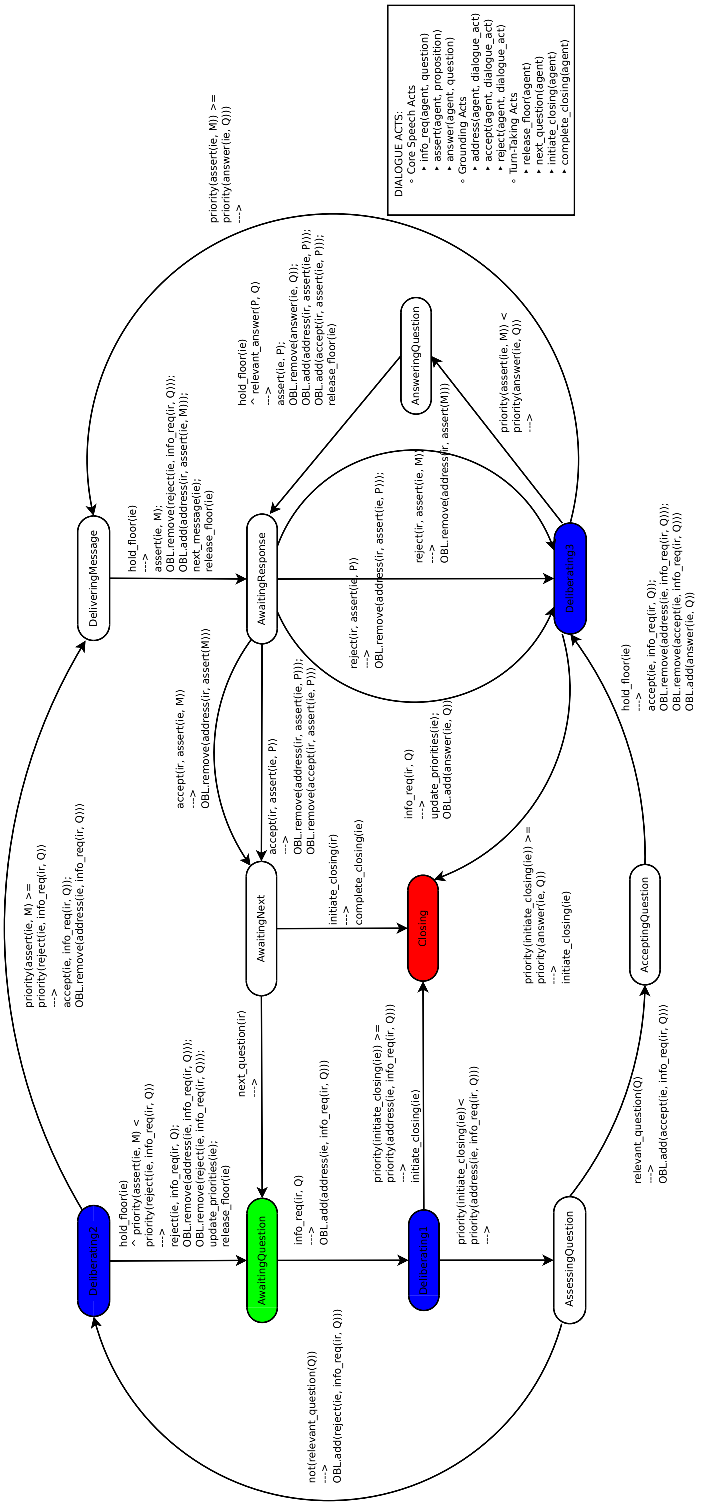


Figure 5.7: A finite state machine specifying the interviewee prototype

assumptions. By allowing potential mismatches in the understanding of utterances it would be possible to explore the relation between grounding and linguistic cooperation. Considering small differences in the dialogue games or in the conversational domains held by each agent would allow for an exploration of cases in which departures from expected behaviour result from different mindsets and cultural backgrounds rather than from egoistic goals. We will come back to this in the next chapter.

The way in which the components of the agents operate with each other to model dialogue behaviour was illustrated by showing in detail how a hand-crafted political interview could result from a simulation. We also showed how a simple change of parameters can lead to interactions with very different degrees of linguistic cooperation. These exercises use two biased random functions to decide whether agents favour their obligations or their private goals and whether they release the floor after each contribution. Although these design decisions introduced non-determinism, they allowed us to abstract from the reasons why dialogue participants decide one way or another and to focus on the management of their actual behaviours with respect to the expectations specified by the rules of the dialogue game. Incorporating a mechanism by which agents can reason on the desirability of their actions in terms of utilities and negative consequences is left as future work.

We continued with a discussion on the differences and similarities of our approach with related research. Particular attention is paid to the line of research in which dialogue management is modelled on the basis of discourse obligations (Traum and Allen, 1994; Poesio and Traum, 1998; Matheson et al., 2000; Kreutel and Matheson, 2003b) and of information state updates.

The chapter closes with the description of a prototype that incorporates some of the elements in the conversational agents. The implementation of the system informed the research and helped devising the approach presented in the thesis. A complete implementation of the model as described above is left as future work. It should be a straightforward task, given the level of detail in the formalisation of each module. A natural next step would then be to generate a large set of dialogues which could be automatically analysed by applying the measure described in Chapter 4 to check that the degree of cooperation effectively grows as we increase the value of the runtime parameters that control the extent to which each agent favours discourse obligations over private goals.

Chapter 6

Conclusions and Future Work

This thesis addresses a range of phenomena in natural language dialogue that we referred to as linguistic non-cooperation. Owing to the assumption that dialogue participants cooperate fully in order to complete a task by means of linguistic interaction, these phenomena are often neglected in dialogue research. We identify and characterise them formally by means of a global dialogue game (i.e a set of rules for introducing and discharging discourse obligations), which captures the behaviour conventionally expected from each participant for a given type of conversation as adopted by a community of language users. Linguistic non-cooperation is defined to occur when participants break these rules by failing to discharge their discourse obligations. We propose a semi-automatic empirical measure of the extent to which the actions of each participant are within the rules of the dialogue game – that is, the degree of linguistic cooperation of a speaker with respect to a conversational setting. This measure is instantiated for a specific dialogue type, political interviews, and evaluated by means of a corpus study.

We assess the reliability and validity of the measure, by testing, respectively, inter-annotator agreement and correlation with human judgement. Finally, we propose an architecture for conversational agents that incorporate these concepts and mechanisms as part of their dialogue management model. We show how this model allows for the simulation of a dialogues in which the degree of cooperation of the participants varied notably, by controlling how often the agents favoured their private goals versus their obligations.

6.1 Original Contributions of the Thesis

This thesis makes the following original contributions to knowledge:

- A definition of cooperative and non-cooperative linguistic behaviour in dialogue, which combines the notions of discourse obligations and dialogue games to specify appropriate behaviour, and allows for the detection of inappropriate actions. The definition is formalised and fully specified for the political interview conversational setting.

- A coding scheme for the manual segmentation, annotation and classification of linguistic behaviour in political interviews. The coding scheme is supported by domain-independent tools, and evaluated for reliability on a corpus of political interviews.

- A domain-independent, automatic method for measuring non-cooperative linguistic behaviour empirically in annotated dialogue. The method is fully implemented, and evaluated for validity on a corpus of political interviews.

- A domain-independent, formal and implementable model of conversational agents that incorporates the concepts and mechanisms above, combining them with the other elements and functions involved in conversation. We show how this addition allows for the generation of a wider range of dialogues, by manipulating parameters that control how agents weigh discourse obligations and private goals when deciding on their contributions.

Let us see where these contributions stand with respect to the research question stated in Chapter 1:

RQ: What elements are needed in a computational model of conversational agents so that they can exhibit and cope with non-cooperative as well as cooperative linguistic behaviour in dialogue, in particular in the domain of political interviews?

The answer to this question, in the light of the results presented in previous chapters, consists of the following elements:

- A set of **dialogue acts** which are the conversational actions that agents can perform and interpret, such as asking questions, making statements, etc. This is a basic requirement both for cooperative and non-cooperative dialogue.

- A **global dialogue game**: the set of rules that establishes how discourse obligations resulting from the performance of dialogue acts are introduced and discharged throughout the conversation. In particular for political interviews these rules state that interviewers are bound to asking questions and politicians to providing answers. Questions should be neutral, relevant and reasonable, while answers should provide all the information requested, and so forth.

- A specification of the agents' **conversational roles**, including their agendas of private goals. In political interviews the roles are the interviewer and the interviewee. Respectively, their goals include asking a specific set of questions and delivering a specific set of key messages.

- A structure, the **information state**, for maintaining a record of the dialogue history, pending obligations and private goals, and a set of **update rules** specifying how the state changes after the performance of each dialogue act.

- A **cooperation measuring mechanism** for interpreting dialogue acts – their own and those of the other party – in terms of how appropriate they are with respect to the dialogue game. In political interviews this mechanism will allow the agent to establish whether a question is adequate or not, whether a statement constitutes a valid reply to a question, etc.

- A **deliberation mechanism** allowing the agents to decide whether to discharge an obligation or favour a private goal at the moment of producing their contributions.

6.2 Future Work

The research presented in this thesis opens up several lines of future work, from extensions to the model to address aspects of dialogue we have not dealt with, to a complete implementation of the agents, to the application of the approach to other conversational domains, to practical – and even commercial – exploitation.

6.2.1 Extensions to the Model

We have made a few simplifying assumptions in order to scope the research to a project of manageable size. Basing the analysis on dialogue transcripts means that we have not dealt with prosody, gestures and other multi-modal aspects of dialogue interaction, as well as sub-utterance elements such as interruptions, incomplete and overlapped speech, etc. Also, the assumption that generation and interpretation are flawless and the omission of clarification subdialogues allowed us to avoid dealing with grounding issues and their connection with linguistic non-cooperation. The relaxation of each of these assumptions would lead to extensions of the model, using the same approach for dealing with non-cooperation: identifying a set of rules that capture expected behaviour at these levels of interaction, and detecting instances in which these rules are broken.

A further line of exploration in this regard would be towards the automation of the empirical measure. Data-driven techniques using machine learning could potentially be used to automatically annotate the dialogues with the labels needed to assess the degree of cooperation. Further, we speculate that the rules of the dialogue game could be learned from a sufficiently large corpus of interviews that are deemed conventional.

6.2.2 Implementation of Conversational Agents

The implementation described at the end of the previous chapter is incomplete and worked more as an aid to the research than as a testbed on which to evaluate the feasibility of the model. However, the level of detail at which the components of the agents have been described should allow for a rather straightforward implementation. A complete implementation would involve the development of one or more topical domains for the agents to talk about.

The use of ontologies for knowledge representation, and of robust, state-of-the-art natural language generation and understanding, would result in more realistic interactions. This would in turn lead to the possibility of evaluating the approach in practice, for instance by having human judges score the resulting interactions between two artificial agents, and also between an agent and a human interlocutor.

6.2.3 Applications of the Approach to other Domains

Although presented as generally domain-independent, the approach has been proposed and evaluated in great detail for the domain of political interviews. Similar applications to other conversational domains in which it is possible to identify a set of rules of expected interaction would be most beneficial in further assessing the suitability of the approach. Examples of such domains include courtroom interrogations, tutoring sessions, doctor-patient exchanges, customer care and many more.

6.2.4 Practical and Commercial Exploitation

The findings reported in this thesis can have a significant impact in real-world scenarios. Practical applications include the semi-automatic – and perhaps, eventually, automatic – analysis of political discourse in terms of speaker compliance[1], and the creation of virtual characters that interviewers and politicians could use to train their interviewing and public speaking skills. In other domains, including these mechanisms in existing dialogue

[1] This also applies to other domains in which it would be useful to automatically or semi-automatically assess the behaviour of participants. In courtroom interaction, for instance, the detection of violations of interrogation rules by either party would be of assistance to judges in charge of keeping order. In customer care, the possibility of detecting deviation from the expected conversation flow, or even instances of abuse and altercation, would contribute in improving the quality of the service and prevent damages to the image of the company.

systems would lead to more natural and robust tutoring, personal assistant and customer care agents. This prospect is appealing not only from the point of view of academic research but also in commercial settings, potentially increasing the quality of available solutions and leading to novel products or services.

6.3 Concluding Remarks

Looking for an answer to the research question behind this thesis has shed light on several areas of research in computational linguistics and political dialogue. We revisited the works of conversational analysts, social psychologists and political scientists, providing a better understanding of the actions of interviewers and politicians during an interview, of the consequences these actions have on the dynamics of the dialogue, and of how all this is perceived by the audience. The accuracy of our model of non-cooperative linguistic behaviour has implications beyond political interviews, and suggests that the approach can be usefully applied to other conversational domains. In the area of dialogue systems, a generalization of our results would lead to the development of systems that can both perform and respond to non-cooperative conversational behaviour, and that can act non-cooperatively, resulting in increased flexibility, robustness and closeness to human behaviour. Finally, as for general knowledge, this research provides a better understanding of dialogue structure and pragmatics by looking at phenomena that have not been addressed before. It is our hope that this work will attract the attention of an increasingly large number of researchers in politics, linguistics, the computational modelling of conversation and beyond.

Si el honor y la sabiduría y la felicidad no son para mí, que sean para otros. Que el cielo exista, aunque mi lugar sea el infierno. Que yo sea ultrajado y aniquilado, pero que en un instante, en un ser, Tu enorme Biblioteca se justifique.

La Biblioteca de Babel
JORGE LUIS BORGES

Bibliography

(Afantenos et al., 2012) Stergos Afantenos, Nicholas Asher, Farah Benamara, Anais Cadilhac, Cedric Dégremont, Pascal Denis, Markus Guhe, Simon Keizer, Alex Lascarides, Oliver Lemon, Philippe Muller, Soumya Paul, Vladimir Popescu, Verena Rieser, and Laure Vieu. 2012. Modelling Strategic Conversation: model, annotation design and corpus. In *Proceedings of SemDial 2012 (SeineDial), 16th Workshop on the Semantics and Pragmatics of Dialogue*, Paris, France, September. Cited on pages 27 and 106.

(Aist, 1998) Gregory Aist. 1998. Expanding a time-sensitive conversational architecture for turn-taking to handle content-driven interruption. In *Proceedings of ICSLP 98 Fifth International Conference on Spoken Language Processing*, pages 413–417. Cited on page 69.

(Alexandersson et al., 1997) Jan Alexandersson, N. Reithinger, and E. Maier. 1997. Insights into the Dialogue Processing of VERBMOBIL. In *Proceedings of the fifth conference on Applied natural language processing*, pages 33–40. Association for Computational Linguistics Morristown, NJ, USA. Cited on page 32.

(Allen and Core, 1997) James Allen and Mark Core. 1997. Draft of DAMSL: Dialog Act Markup in Several Layers. Technical report, University of Rochester. Cited on pages 69, 105, 110, and 111.

(Allen and Perrault, 1980) James Allen and Raymond Perrault. 1980. Analyzing Intentions in Utterances. *Artificial Intelligence*, 15:143–178. Cited on page 33.

(Allen and Schubert, 1991) James Allen and L.K. Schubert. 1991. The TRAINS project. TRAINS Technical Note 91-1. *Computer Science Dept. University of Rochester*. Cited on pages 32 and 35.

(Allen, 1995) James Allen. 1995. *Natural language understanding*. Benjamin-Cummings Publishing Co., Inc. Redwood City, CA, USA. Cited on pages 33 and 222.

(Allwood et al., 2000) Jens Allwood, David Traum, and Kristiina Jokinen. 2000. Cooperation, dialogue and ethics. *International Journal of Human-Computer Studies*, 53(6):871–914. Cited on page 29.

(Allwood, 1976) Jens Allwood. 1976. Linguistic Communication as Action and Cooperation. *Gothenburg monographs in linguistics*, 2:637–663. Cited on page 29.

(Anderson et al., 1991) Anne H Anderson, Miles Bader, Ellen Gurman Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline C Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, and Henry S. Thompson. 1991. The HCRC Map Task Corpus. *Language and speech*, 34(4):351–366. Cited on page 106.

(Artstein and Poesio, 2008) Ron Artstein and Massimo Poesio. 2008. Intercoder Agreement for Computational Liguistics. *Computational Linguistics*, 34(4):555–596. Cited on pages 108, 143, 144, 145, 146, 147, 148, and 149.

(Asher and Lascarides, 2003) Nicholas Asher and Alex Lascarides. 2003. *Logics of conversation*. Cambridge University Press. Cited on pages 24 and 25.

(Asher and Lascarides, 2008) Nicholas Asher and Alex Lascarides. 2008. Making the right commitments in dialogue. In *Proceedings of the Fall 2008 Workshop in Philosophy and Linguistics*. University of Michigan. Cited on pages 24 and 26.

(Asher and Lascarides, in press) Nicholas Asher and Alex Lascarides. in press. Strategic Conversation. To appear in *Semantics and Pragmatics*. Cited on pages 24, 25, 26, 27, 31, 71, and 72.

(Asher et al., 2012) Nicholas Asher, Alex Lascarides, Oliver Lemon, Markus Guhe, Verena Rieser, Philippe Muller, Stergos Afantenos, Farah Benamara, Laure Vieu, Pascal Denis, Soumya Paul, Simon Keizer, and Cedric Dégremont. 2012. Modelling Strategic Conversation: the STAC project. In *Proceedings of SemDial 2012 (SeineDial), 16th Workshop on the Semantics and Pragmatics of Dialogue*, Paris, France, September. Cited on pages 27, 72, and 106.

(Asher, 2012) Nicholas Asher. 2012. The non cooperative basis of implicatures. In Denis Bchet and Alexander Dikovsky, editors, *Logical Aspects of Computational Linguistics*, volume 7351 of *Lecture Notes in*

Computer Science, pages 45–57. Springer Berlin Heidelberg. Cited on page 71.

(Attardo, 1997) Salvatore Attardo. 1997. Locutionary and perlocutionary cooperation: The perlocutionary cooperative principle. *Journal of Pragmatics*, 27(6):753–779. Cited on pages 8, 25, and 61.

(Austin, 1962) J.L. Austin. 1962. *How to do Things with Words*. Harvard University William James Lectures 1955. Clarendon Press, Oxford. Cited on page 33.

(Bazzanella, 2011) Carla Bazzanella. 2011. Redundancy, repetition, and intensity in discourse. *Language Sciences*, 33(2):243–254. Cited on page 87.

(Benotti, 2009) Luciana Benotti. 2009. Clarification Potential of Instructions. In *Proceedings of the 2009 SIGDIAL Conference on Discourse and Dialogue*, pages 196–205, London, United Kingdom, September. Cited on pages 71 and 72.

(Benotti, 2010) Luciana Benotti. 2010. *Implicature as an Interactive Process*. Ph.D. thesis, Ph. D. thesis, Université Henri Poincaré, INRIA Nancy Grand Est, France. Supervised by P. Blackburn. Reviewed by N. Asher and B. Geurts. Cited on page 71.

(Beun, 2001) Robbert-Jan Beun. 2001. On the generation of coherent dialogue: A computational approach. *Pragmatics and Cognition*, 9(1):37–68. Cited on pages 44 and 267.

(Bird et al., 2009) Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python*. O'Reilly Media, Incorporated. Cited on page 149.

(Blaylock and Allen, 2005) Nathan Blaylock and James Allen. 2005. A collaborative problem-solving model of dialogue. In *Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue*, pages 200–211, Lisbon, Portugal. Cited on page 1.

(Boella et al., 1999) Guido Boella, Rossana Damiano, Leonardo Lesmo, and Liliana Ardissono. 1999. Conversational cooperation: the leading role of intentions. In *Proceedings of Amstelogue 99, the 3rd Workshop on the Semantics and Pragmatics of Dialogue*, May. Cited on pages 41, 61, and 75.

(Bos et al., 2003) Johan Bos, Ewan Klein, Oliver Lemon, and Tetsushi Oka. 2003. Dipper: Description and formalisation of an information-state update dialogue system architecture. In *4th SIGdial Workshop on Discourse and Dialogue*, pages 115–124. Cited on page 269.

(Brown-Schmidt, 2012) Sarah Brown-Schmidt. 2012. Beyond common and privileged: Gradient representations of common ground in real-time language use. *Language and Cognitive Processes*, 27(1):62–89. Cited on page 71.

(Bull and Mayer, 1993) Peter Bull and K. Mayer. 1993. How not to answer questions in political interviews. *Political Psychology*, pages 651–666. Cited on pages 53, 88, 89, 106, 110, and 126.

(Bull, 1994) Peter Bull. 1994. On identifying questions, replies, and non-replies in political interviews. *Journal of Language and Social Psychology*, 13(2):115. Cited on pages 53, 106, 110, 112, and 126.

(Bull, 2003) Peter Bull. 2003. *The Microanalysis of Political Communication: Claptrap and ambiguity*. Routledge. Cited on pages 53, 89, 106, 110, 112, and 126.

(Bull, 2008) Peter Bull. 2008. Slipperiness, Evasion, and Ambiguity. *Journal of Language and Social Psychology*, 27(4):333. Cited on page 53.

(Bunt et al., 2010) Harry Bunt, Jan Alexandersson, Jean Carletta, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Kiyong Lee, Volha Petukhova, Andrei Popescu-Belis, Laurent Romary, Claudia Soria, and David Traum. 2010. Towards an ISO Standard for Dialogue Act Annotation. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA). Cited on pages 106, 110, and 155.

(Bunt et al., 2012) Harry Bunt, Jan Alexandersson, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Volha Petukhova, Andrei Belis-Popescu, and David Traum. 2012. ISO 24617-2: A semantically-based standard for dialogue annotation. In *LREC 2012*, Istanbul, Turkey. Cited on pages 105, 106, 110, 111, 155, and 267.

(Bunt, 1994) Harry Bunt. 1994. Context and Dialogue Control. *THINK Quarterly*, 3. Cited on pages 8, 30, 32, 33, 34, and 61.

(Bunt, 2009) Harry Bunt. 2009. The DIT++ taxonomy for functional dialogue markup. In *AAMAS 2009 Workshop, Towards a Standard Markup Language for Embodied Dialogue Acts*, pages 13–24. Cited on pages 105, 106, 110, and 155.

(Bunt, 2011) Harry Bunt. 2011. The semantics of dialogue acts. In *Proceedings of the Ninth International Conference on Computational Semantics*, pages 1–13, Stroudsburg, PA, USA. Association for Computational Linguistics. Cited on page 105.

(Bunt, 2012) Harry Bunt. 2012. The semantics of feedback. In *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL 2012)*, pages 118–127, Paris, France. University Paris-Diderot, Paris Sorbonne-Cite. Cited on page 72.

(Buschmeier and Kopp, 2012) Hendrik Buschmeier and Stefan Kopp. 2012. Understanding how well you understood – Context-sensitive interpretation of multimodal user feedback. In *Proceedings of the 12th International Conference on Intelligent Virtual Agents*, pages 517–519, Santa Cruz, CA. Cited on page 72.

(Carletta et al., 1997) Jean Carletta, Stephen Isard, G. Doherty-Sneddon, Amy Isard, J. C. Kowtko, and A. H. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13–31. Cited on pages 1, 33, 41, 44, 73, 105, 106, 110, 145, 148, 151, and 154.

(Carlson, 1983) L. Carlson. 1983. *Dialogue games: An approach to discourse analysis*. Kluwer Academic Publishers. Cited on pages 44 and 73.

(Cassell, 2001) Justine Cassell. 2001. Embodied Conversational Agents: Representation and Intelligence in User Interfaces. *AI Magazine*, 22(4):67–84. Cited on page 46.

(Cavicchio, 2010) Federica Cavicchio. 2010. *Computational Modeling of (un)Cooperation: The Role of Emotions*. Ph.D. thesis, University of Trento. CIMeC. Cited on pages 105, 106, and 107.

(Chu-Carroll and Carberry, 1998) J. Chu-Carroll and Sandra Carberry. 1998. Collaborative response generation in planning dialogues. *Computational Linguistics*, 24(3):355–400. Cited on pages 1 and 33.

(Clark and Brennan, 1991) Herbert Clark and S.E. Brennan. 1991. Grounding in communication. *Perspectives on socially shared cognition*, pages 127–149. Cited on pages 33 and 71.

(Clark and Schaefer, 1989) Herbert Clark and Edward Schaefer. 1989. Contributing to Discourse. *Cognitive science*, 13(2):259–294. Cited on pages 1 and 33.

(Clark, 1996) Herbert Clark. 1996. *Using Language*. Cambridge University Press, Cambridge, MA. Cited on pages 9, 23, 26, 31, 268, and 270.

(Clayman and Heritage, 2002) Steven Clayman and John Heritage. 2002. *The News Interview: Journalists and Public Figures on the Air*. Cambridge University Press. Cited on pages 49, 50, and 64.

(Clayman, 1988) Steven E. Clayman. 1988. Displaying neutrality in television news interviews. *Soc. Probs.*, 35:474. Cited on pages 48, 52, and 64.

(Cohen and Levesque, 1990) Philip Cohen and Hector Levesque. 1990. Rational interaction as the basis for communication. *Intentions in communication*, pages 221–256. Cited on page 32.

(Cohen and Levesque, 1991) Philip Cohen and Hector Levesque. 1991. Confirmations and joint action. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 951–957, Sydney, Australia. Cited on pages 1, 35, and 74.

(Cohen, 1960) Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurment*, 20(1):37–46. Cited on pages 102, 141, and 142.

(Cooper and Larsson, 1999) Robin Cooper and Staffan Larsson. 1999. Dialogue moves and information states. *Proc. of the Third IWCS, Tilburg.* Cited on page 219.

(Cooper, 1998) Robin Cooper. 1998. Mixing situation theory and type theory to formalize information states in dialogue exchanges. In *Proceedings of TwenDial, the thirteenth Twente Workshop on Language Technology*, Twente, Netherlands. Cited on page 219.

(Cooper, 2013) Robin Cooper. 2013. Clarification and generalized quantifiers. *Dialogue {&* Discourse Cited on page 72.

, 4(1).

(Crook et al., 2010) Nigel Crook, Cameron Smith, Marc Cavazza, Stephen Pulman, Roger Moore, and Johan Boye. 2010. Handling user interruptions in an embodied conversational agent. In *Proceedings of the AAMAS International Workshop on Interacting with ECAs as Virtual Characters, Toronto*, pages 27–33. Cited on page 71.

(Davies and Fleiss, 1982) Mark Davies and Joseph L. Fleiss. 1982. Measuring agreement for multinomial data. *Biometrics*, pages 1047–1051. Cited on pages 102, 141, and 142.

(Davies, 1994) Bethan L. Davies. 1994. To cooperate or not to cooperate - is that the question? In *Proceedings of the Edinburgh Linguistics Department Conference*, pages 17–32. Citeseer. Cited on page 106.

(Davies, 1997) Bethan L. Davies. 1997. *An Empirical Examination of Cooperation, Effort and Risk in Task-oriented Dialogue.* Ph.D. thesis, University of Edinburgh. Cited on pages 105 and 106.

(Davies, 2006) Bethan L. Davies. 2006. Testing dialogue principles in task-oriented dialogues: An exploration of cooperation, collaboration, effort and risk. *Leeds Working Papers in Linguistics and Phonetics*, 11:30–64. Cited on pages 106 and 107.

(DeVault et al., 2008) David DeVault, David Traum, and Ron Artstein. 2008. Making grammar-based generation easier to deploy in dialogue systems. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, pages 198–207. Association for Computational Linguistics. Cited on page 222.

(DeVault et al., 2009) David DeVault, Kenji Sagae, and David Traum. 2009. Can i finish?: learning when to respond to incremental interpretation results in interactive dialogue. In *Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, SIGDIAL '09, pages 11–20, Stroudsburg, PA, USA. Association for Computational Linguistics. Cited on page 71.

(DeVault et al., 2011) David DeVault, Kenji Sagae, and David Traum. 2011. Incremental interpretation and prediction of utterance meaning for interactive dialogue. *Dialogue {&* Discourse Cited on pages 71 and 222.

, 2(1):143–170.

(Di Eugenio and Glass, 2004) Barbara Di Eugenio and Michael Glass. 2004. The kappa statistic: A second look. *Computational Linguistics*, 30(1):95–101. Cited on page 145.

(Dzikovska et al., 2010) Myroslava O Dzikovska, Johanna Moore, Natalie Steinhauser, and Gwendolyn Campbell. 2010. The impact of interpretation problems on tutorial dialogue. In *Proceedings of the ACL 2010 Conference Short Papers*, pages 43–48. Association for Computational Linguistics. Cited on page 72.

(Dzikovska et al., 2012) Myroslava O Dzikovska, Rodney D. Nielsen, and Chris Brew. 2012. Towards effective tutorial feedback for explanation questions: a dataset and baselines. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 200–210, Stroudsburg, PA, USA. Association for Computational Linguistics. Cited on page 72.

(Elliott and Bull, 1996) Judy Elliott and Peter Bull. 1996. A question of threat: face threats in questions posed during televised political interviews. *Journal of community {&* applied social psychology Cited on page 53.

, 6(1):49–72.

(Fikes and Nilsson, 1972) Richard Fikes and Nils Nilsson. 1972. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2(3):189–208. Cited on page 220.

(Finlayson, 2001) Alan Finlayson. 2001. The problem of the political interview. *The Political Quarterly*, 72(3):335–344. Cited on page 53.

(Fleiss, 1971) Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378. Cited on pages 102, 141, and 142.

(Frederking, 1996) R. Frederking. 1996. Grice's maxims: do the right thing. *Proc. of AAAI SpringSymp. on Compl. Implicature: Computational Approaches to Interpreting and Generating Conversational Implicature*. Cited on page 23.

(Galliers, 1988) Julia Rose Galliers. 1988. *A theoretical framework for computer models of cooperative dialogue, acknowledging multiagent conflict*. Ph.D. thesis, The Open University. Cited on page 37.

(Georgila et al., 2011) Kallirroi Georgila, Ron Artstein, Angela Nazarian, Michael Rushforth, David Traum, and Katia Sycara. 2011. An annotation scheme for cross-cultural argumentation and persuasion dialogues. In *Proceedings of the SIGDIAL 2011 Conference*, pages 272–278. Association for Computational Linguistics. Cited on pages 72 and 105.

(Ginzburg, 1996) Jonathan Ginzburg. 1996. Interrogatives: Questions, facts and dialogue. *The handbook of contemporary semantic theory*, 5:359–423. Cited on pages 6, 41, and 267.

(Ginzburg, 1997) Jonathan Ginzburg. 1997. Querying and Assertion in Dialogue. Book chapter draft. Cited on pages 41, 267, and 268.

(Ginzburg, 2012) Jonathan Ginzburg. 2012. *The Interactive Stance*. Oxford University Press. Cited on pages 44, 72, 267, and 268.

(Goffman, 1979) Erving Goffman. 1979. Footing. *Semiotica*, 25(1-2):1–30. Cited on page 51.

(Greatbatch, 1986) David Greatbatch. 1986. Aspects of topical organization in news interviews: the use of agenda-shifting procedures by interviewees. *Media, Culture & Society*, 8(4):441–455. Cited on pages 48 and 52.

(Greatbatch, 1988) David Greatbatch. 1988. A turn-taking system for British news interviews. *Language in Society*, 17(03):401–430. Cited on pages 48, 49, 64, and 95.

(Grice, 1975) H. P. Grice. 1975. Logic and conversation. *Syntax and Semantics*, 3:41–58. Cited on pages 1, 21, 22, 23, 24, 25, and 107.

(Grosz and Sidner, 1986) B.J. Grosz and C.L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204. Cited on pages 1 and 41.

(Grosz and Sidner, 1990) B.J. Grosz and C.L. Sidner. 1990. Plans for discourse. *Intentions in communication*, pages 417–444. Cited on pages 1, 33, 35, and 74.

(Hamblin, 1970) C.L. Hamblin. 1970. *Fallacies*. Methuen. Cited on pages 43, 73, and 75.

(Harris, 1991) S. Harris. 1991. Evasive action: How politicians respond to questions in political interviews. *Broadcast talk*, 7699. Cited on page 53.

(Heritage and Greatbatch, 1991) John Heritage and David Greatbatch. 1991. On the institutional character of institutional talk: The case of news interviews. *Talk and social structure*, pages 93–137. Cited on pages 49, 50, 52, and 64.

(Heritage, 1985) John Heritage. 1985. Analyzing news interviews: Aspects of the production of talk for an overhearing audience. *Handbook of discourse analysis*, 3:95–117. Cited on pages 48 and 95.

(Heritage, 1998) John Heritage. 1998. Conversation analysis and institutional talk. Analyzing distinctive turn-taking systems. In *Proceedings of the 6th International Congress of IADA (International Association for Dialog Analysis), Tubingen, Niemeyer*. Cited on pages 5, 51, and 64.

(Heritage, 2005) John Heritage. 2005. Conversation analysis and institutional talk. *Handbook of language and social interaction*, pages 103–147. Cited on page 64.

(Hutchby, 2011) I. Hutchby. 2011. Non-neutrality and argument in the hybrid political interview. *Discourse Studies*, 13(3):349. Cited on page 87.

(Initiative, 1997) Discourse Resource Initiative. 1997. Standards for dialogue coding in natural language processing. In *TechnicalReport167, Dagstuhl-Seminar*. Cited on pages 39, 69, 110, and 267.

(Jameson and Weis, 1995) Anthony Jameson and Thomas Weis. 1995. How to juggle discourse obligations. In *Proceedings of the Symposium on Conceptual and Semantic Knowledge in Language Generation*, pages 171–185. Cited on page 75.

(Jameson et al., 1994) A. Jameson, B. Kipper, A. Ndiaye, R. Schaefer, J. Simons, T. Weis, and D. Zimmermann. 1994. Cooperating to be Noncooperative: The Dialog System PRACMA. *Lecture Notes in Computer Science*, pages 106–106. Cited on page 45.

(Jameson, 1989) A. Jameson. 1989. But what will the listener think? Belief ascription and image maintenance in dialog. *User Models in Dialog Systems. Springer-Verlag*, pages 255–312. Cited on page 45.

(Jekat et al., 1995) S. Jekat, A. Klein, E. Maier, I. Maleck, M. Mast, and J. Quantz. 1995. Dialogue acts in VERBMOBIL. *Verbmobil Report*, 65. Cited on page 32.

(Jurafsky et al., 1997) Daniel Jurafsky, Elizabeth Shriberg, and Debra Biasca. 1997. Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual, Draft 13. *University of Colorado, Boulder Institute of Cognitive Science Technical Report*, pages 97–02. Cited on page 154.

(Karunatillake et al., 2009) Nishan C Karunatillake, Nicholas R Jennings, Iyad Rahwan, and Peter McBurney. 2009. Dialogue games that agents play within a society. *Artificial intelligence*, 173(9):935–981. Cited on pages 43 and 74.

(Keenan, 1976) E.O. Keenan. 1976. The universality of conversational postulates. *Language in Society*, pages 67–80. Cited on page 23.

(Kiefer, 1979) F. Kiefer. 1979. What do conversational maxims explain? *Linguistical Investigationes. Revue Internationale de Linguistique Française et de Linguistique Générale Paris*, 3(1):57–74. Cited on page 23.

(Kowtko et al., 1992) J. C. Kowtko, Stephen Isard, and G.M. Doherty. 1992. Conversational games within dialogue. Research Paper HCRC/RP-31. *Human Communication Research Centre, University of Edinburgh.* Cited on pages 44, 73, and 74.

(Kreutel and Matheson, 1999) Jörn Kreutel and Colin Matheson. 1999. Modelling Questions and Assertions in Dialogue Using Obligations. In *Proceedings of Amstelog 99, the 3rd Workshop on the Sematics and Pragmatics of Dialogue.* Cited on pages 38, 41, and 268.

(Kreutel and Matheson, 2000) Jörn Kreutel and Colin Matheson. 2000. Information states, obligations and intentional structure in dialogue modelling. In *Proceedings of the 3rd International Workshop on Human-Computer Conversation.* Citeseer. Cited on pages 38, 41, 75, and 268.

(Kreutel and Matheson, 2001) Jörn Kreutel and Colin Matheson. 2001. Co-operation and strategic acting in discussion scenarios. In *Proceedings of the Workshop on Coordination and Action*. Citeseer. Cited on pages 38, 41, and 48.

(Kreutel and Matheson, 2003a) Jörn Kreutel and Colin Matheson. 2003a. Context-Dependent Interpretation and Implicit Dialogue Acts. In *Perspectives on Dialogue in the New Millenium*, pages 179–192. John Benjamins. Cited on pages 41 and 95.

(Kreutel and Matheson, 2003b) Jörn Kreutel and Colin Matheson. 2003b. Incremental information state updates in an obligation-driven dialogue model. *Logic Journal of IGPL*, 11(4):485–511. Cited on pages 38, 41, 48, 75, 267, 268, 271, and 277.

(Krippendorff, 1980) Klaus Krippendorff. 1980. *Content Analysis: An Introduction to Its Methodology*. Sage Publications, Incorporated. Cited on page 145.

(Krippendorff, 1995) Klaus Krippendorff. 1995. *On the reliability of unitizing continuous data*. Sociological Methodology. Cited on pages 102, 141, 146, and 147.

(Krippendorff, 2003) Klaus Krippendorff. 2003. *Content Analysis: An Introduction to Its Methodology*. Sage Publications, Incorporated, second edition edition, December. Cited on page 141.

(Krippendorff, 2004) Klaus Krippendorff. 2004. Measuring the Reliability of Qualitative Text Analysis Data. *Quality & Quantity*, 38(6):787–800, December. Cited on pages 102, 141, 142, and 146.

(Landis and Koch, 1977) J.R. Landis and G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics*, pages 159–174. Cited on page 145.

(Larsson and Traum, 2000) Staffan Larsson and David Traum. 2000. Information state and dialogue management in the TRINDI dialogue move engine toolkit. *Natural Language Engineering*, 6(3-4):323–340. Cited on pages 40, 267, and 272.

(Larsson et al., 2000) Staffan Larsson, A. Berman, Johan Bos, L Gronqvist, P. Ljunglöf, and David Traum. 2000. TrindiKit 2.0 Manual. *Deliverable D5.3, TRINDI Project*. Cited on pages 40 and 272.

(Levin and Moore, 1977) James A. Levin and James A. Moore. 1977. Dialogue-games: Metacommunication structures for natural language interaction. *Cognitive Science*, 1(4):395 – 420. Cited on pages 44 and 73.

(Levinson, 1979) Stephen C Levinson. 1979. Activity types and language. *Linguistics*, 17(5-6):365–400. Cited on pages 42 and 43.

(Levinson, 1983) Stephen C Levinson. 1983. *Pragmatics*. Cambridge University Press. Cited on pages 12 and 24.

(Lewin, 2000) Ian Lewin. 2000. A formal model of Conversational Game Theory. In *In Proc. Gotalog-00, 4 th Workshop on the Semantics and Pragmatics of Dialogue, Gothenburg*. Cited on pages 44 and 73.

(Litman et al., 2010) Diane Litman, Johanna Moore, Myroslava Dzikovska, and Elaine Farrow. 2010. Using Natural Language Processing to Analyze Tutorial Dialogue Corpora Across Domains Modalities. In *Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*, pages 149–156, Amsterdam, The Netherlands, The Netherlands. IOS Press. Cited on page 72.

(Ljunglöf, 2009) P. Ljunglöf. 2009. trindikit.py: An open-source Python library for developing ISU-based dialogue systems. In *IWSDS'09, 1st International Workshop on Spoken Dialogue Systems Technology*. Cited on page 272.

(Lu et al., 2007) Xin Lu, Barbara Di Eugenio, Trina C. Kershaw, Stellan Ohlsson, and Andrew Corrigan-Halpern. 2007. Expert vs. Non-expert Tutoring: Dialogue Moves, Interaction Patterns and Multi-utterance Turns. In *Computational Linguistics and Intelligent Text Processing*, pages 456–467. Springer Berlin Heidelberg. Cited on page 72.

(Mann, 1988) William C. Mann. 1988. Dialogue games: Conventions of human interaction. *Argumentation*, 2(4):511–532. Cited on pages 44 and 73.

(Mann, 2002) William C Mann. 2002. Dialogue macrogame theory. In *Proceedings of the 3rd SIGdial workshop on Discourse and Dialogue*, pages 129–141. Association for Computational Linguistics. Cited on pages 44 and 73.

(Matheson et al., 2000) Colin Matheson, Massimo Poesio, and David Traum. 2000. Modelling grounding and discourse obligations using update rules. In *Proceedings of the 1st conference on North American chapter of the Association for Computational Linguistics*, pages 1–8, Morgan Kaufmann Publishers Inc. Morgan Kaufmann Publishers Inc. Cited on pages 38, 40, 48, 75, 268, 271, and 277.

(Maudet and Moore, 2001) Nicolas Maudet and David Moore. 2001. Dialogue games as dialogue models for interacting with, and via, computers. *Informal Logic*, 21(3). Cited on pages 43 and 74.

(McBurney and Parsons, 2009) Peter McBurney and Simon Parsons. 2009. Dialogue games for agent argumentation. In *Argumentation in artificial intelligence*, pages 261–280. Springer. Cited on pages 43 and 74.

(McBurney et al., 2003) Peter McBurney, Rogier M Van Eijk, Simon Parsons, and Leila Amgoud. 2003. A dialogue game protocol for agent purchase negotiations. *Autonomous Agents and Multi-Agent Systems*, 7(3):235–273. Cited on page 72.

(Mostow et al., 2003) Jack Mostow, Gregory Aist, Paul Burkhead, A Corbett, Andrew Cuneo, S Eitelman, Cathy Huang, Brian Junker, Mary Beth Sklar, and Brian Tobin. 2003. Evaluation of an automated Reading Tutor that listens: Comparison to human tutoring and classroom instruction. *Journal of Educational Computing Research*, 29(1):61–117, January. Cited on page 69.

(Muskens, 1994) R. Muskens. 1994. A compositional discourse representation theory. In *Proceedings of the 9th Amsterdam Colloquium*, pages 467–486. Cited on page 39.

(Piwek et al., 2007) Paul Piwek, David Hardcastle, and Richard Power. 2007. Dialogue games for crosslingual communication, June. Cited on pages 43 and 74.

(Piwek, 1998) Paul Piwek. 1998. *Logic, Information, and Conversation.* Ph.D. thesis, Eindhoven University of Technology. Cited on pages 43 and 267.

(Piwek, 2006) Paul Piwek. 2006. Perspectives on Dialogue: Introduction to this Special Issue. *Research on Language & Computation*, 4(2):143–152. Cited on page 62.

(Piwek, 2008) Paul Piwek. 2008. Presenting Arguments as Fictive Dialogue. In *Proceedings of 8th Workshop on Computational Models of Natural Argument (in conjunction with ECAI 2008)*, Patras, Greece. Cited on page 72.

(Plüss et al., 2011) Brian Plüss, David DeVault, and David Traum. 2011. Toward Rapid Development of Multi-Party Virtual Human Negotiation Scenarios. In *SemDial 2011: Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue*, November. Cited on pages 47 and 72.

(Plüss, 2009) Brian Plüss. 2009. Towards a Computational Pragmatics for Non-Cooperative Dialogue. Technical report, Department of Computing, The Open University, UK. Cited on page 68.

(Plüss, 2010) Brian Plüss. 2010. Non-cooperation in dialogue. In *ACLstudent '10: Proceedings of the ACL 2010 Student Research Workshop*, pages 1–6, Uppsala, Sweden, July. Association for Computational Linguistics. Cited on pages 68 and 107.

(Poesio and Traum, 1997) Massimo Poesio and David Traum. 1997. Conversational actions and discourse situations. *Computational intelligence*, 13(3):309–347. Cited on pages 37, 38, 39, 69, 110, 267, and 268.

(Poesio and Traum, 1998) Massimo Poesio and David Traum. 1998. Towards an Axiomatization of Dialogue Acts. In *Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues*, pages 207–222. Cited on pages 37, 38, 39, 69, 110, 267, 268, 271, and 277.

(Poesio et al., 1999) Massimo Poesio, Robin Cooper, Staffan Larsson, Colin Matheson, and David Traum. 1999. Annotating conversations for information state updates. In *Proceedings of Amstelogue*. Cited on pages 105 and 267.

(Power, 1979) Richard Power. 1979. The organisation of purposeful dialogues. *Linguistics*, 17:107–152. Cited on pages 1, 33, 44, 73, and 74.

(Prince, 1982) Ellen F Prince. 1982. Grice and Universality: a Reappraisal. Technical report, Penn Linguistics Colloquium. Cited on pages 22, 23, and 24.

(Pulman, 1997) Stephen Pulman. 1997. Conversational games, belief revision and bayesian networks. In J. Landsbergen et al., editors, *Proceedings of CLIN VII, the 7th Computational Linguistics in the Netherlands meeting*, pages 1–25. Cited on pages 44 and 73.

(Pulman, 2002) Stephen Pulman. 2002. Relating dialogue games to information state. *Speech Communication*, 36(1-2):15–30. Cited on pages 44 and 73.

(Purver, 2004) Matthew Purver. 2004. *The Theory and Use of Clarification Requests in Dialogue*. Ph.D. thesis, King's College, University of London. Cited on page 72.

(Purver, 2006) Matthew Purver. 2006. CLARIE: Handling Clarification Requests in a Dialogue System. *Research on Language {& Computation* Cited on page 72.

, 4(2-3):259–288.

(R Development Core Team, 2011) R Development Core Team, 2011. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Cited on page 197.

(Reed and Long, 1997) C. Reed and Derek Long. 1997. Collaboration, co-operation and dialogue classification. *Working Notes of the IJCAI97 Workshop on Collaboration, Cooperation and Conflict in Dialogue Systems, IJCAI 97*, pages 73–78. Cited on pages 8, 29, and 61.

(Reiter and Dale, 2000) Ehud Reiter and Robert Dale. 2000. *Building natural language generation systems*. Cambridge university press. Cited on page 222.

(Rieser and Lemon, 2011) Verena Rieser and Oliver Lemon. 2011. *Reinforcement Learning for Adaptive Dialogue Systems: A Data-driven Methodology for Dialogue Management and Natural Language Generation*. Springer-Verlag New York Inc. Cited on pages 71 and 107.

(Rieser and Moore, 2005) Verena Rieser and Johanna Moore. 2005. Implications for generating clarification requests in task-oriented dialogues. In *Proc of ACL*, pages 239–246. Cited on page 72.

(Roque and Traum, 2007) A. Roque and David Traum. 2007. A model of compliance and emotion for potentially adversarial dialogue agents. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*. Cited on page 47.

(Roque and Traum, 2008) Antonio Roque and David Traum. 2008. Degrees of grounding based on evidence of understanding. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, pages 54–63. Association for Computational Linguistics. Cited on page 71.

(Sacks et al., 1974) H. Sacks, Emanuel Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language*, pages 696–735. Cited on pages 9, 33, and 51.

(Sadek et al., 1996) M.D. Sadek, A. Ferrieux, A. Cozannet, P. Bretier, F. Panaget, and J. Simonin. 1996. Effective human-computer cooperative spoken dialogue: the AGSdemonstrator. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*. Cited on page 32.

(Sadek et al., 1997) M.D. Sadek, P. Bretier, and F. Panaget. 1997. ARTIMIS: Natural dialogue meets rational agency. In *international joint conference on artificial intelligence*, pages 1030–1035. Cited on page 32.

(Sadek, 1992) M.D. Sadek. 1992. A study in the logic of intention. In *Proceedings of the 3rd Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 462–473. Cited on page 32.

(Schegloff, 1968) Emanuel Schegloff. 1968. Sequencing in conversational openings. *American anthropologist*, pages 1075–1095. Cited on page 33.

(Schegloff, 1987) Emanuel Schegloff. 1987. Some Sources of Misunderstanding in Talk-in-Interaction. *Linguistics*, 8:201–218. Cited on page 72.

(Schegloff, 1988) Emanuel A. Schegloff. 1988. From Interview to Confrontation: Observations of the Bush/Rather Encounter. *Research on Language & Social Interaction*, 22(1-4):215–240. Cited on page 51.

(Schlangen and Fernández, 2007) David Schlangen and Raquel Fernández. 2007. Beyond Repair–Testing the Limits of the Conversational Repair System. In *Proceedings of the 2007 SIGDIAL Workshop on Discourse and Dialogue*, pages 51–54, Antwerp, Belgium. Cited on page 72.

(Schlangen and Skantze, 2011) David Schlangen and Gabriel Skantze. 2011. A general, abstract model of incremental dialogue processing. *Dialogue* {*&* Discourse Cited on page 71.

, 2(1):83–111.

(Scott, 1955) William A. Scott. 1955. Reliability of content analysis: The case of nominal scale coding. *Public opinion quarterly*, 19(3):321–325. Cited on pages 102, 141, and 142.

(Searle, 1969) John Searle. 1969. *Speech Acts.* An Essay in the Philosophy of Language. Cambridge University Press. Cited on pages 33 and 267.

(Searle, 1979) John Searle. 1979. A Taxonomy of lllocutionary Acts. *Expression and meaning: studies in the theory of speech acts*, pages 1–29. Cited on pages 10 and 33.

(Siegel and Castellan, 1988) Sidney Siegel and N. John Castellan. 1988. *Nonparametric statistics for the behavioral sciences.* McGraw-Hill, New York, 2 edition, January. Cited on page 142.

(Skantze, 2007) Gabriel Skantze. 2007. *Error Handling in Spoken Dialogue Systems.* Ph.D. thesis, KTH - Royal Institute of Technology, Sweden. Cited on page 72.

(Smith et al., 2011) Cameron Smith, Nigel Crook, Simon Dobnik, Daniel Charlton, Johan Boye, Stephen Pulman, Raul Santos De La Camara, Markku Turunen, David Benyon, Jay Bradley, Björn Gambäck, Preben Hansen, Oli Mival, Nick Webb, and Marc Cavazza. 2011. Interaction Strategies for an Affective Conversational Agent. *Presence: Teleoperators and Virtual Environments*, 20(5):395–411, January. Cited on pages 69 and 71.

(Sperber and Wilson, 1982) D. Sperber and D. Wilson. 1982. Mutual knowledge and relevance in theories of comprehension. *Mutual knowledge*, pages 61–87. Cited on page 23.

(Steedman and Petrick, 2007) Mark Steedman and Ronald Petrick. 2007. Planning Dialog Actions. In *Proceedings of the 2007 SIGDIAL Workshop on Discourse and Dialogue*, pages 265–272, Antwerp, Belgium. Cited on page 220.

(Stoyanchev and Piwek, 2010a) Svetlana Stoyanchev and Paul Piwek. 2010a. Annotation Scheme for Authored Dialogues. Version 1.1. Technical Report 2010/15, Centre for Research in Computing, The Open University, July. Cited on pages 105, 106, and 110.

(Stoyanchev and Piwek, 2010b) Svetlana Stoyanchev and Paul Piwek. 2010b. Constructing the CODA corpus: A parallel corpus of monologues and expository dialogues. In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010)*, Malta, May. Cited on pages 72, 110, 138, 154, and 312.

(Stoyanchev and Piwek, 2010c) Svetlana Stoyanchev and Paul Piwek. 2010c. Harvesting re-usable high-level rules for expository dialogue generation. In *Proceedings of the 6th International Natural Language Generation Conference*, pages 145–154. Association for Computational Linguistics. Cited on page 72.

(Taylor et al., 1996) J. A. Taylor, Jean Carletta, and C. Mellish. 1996. Requirements for belief models in cooperative dialogue. *User Modeling and User-Adapted Interaction*, 6(1):23–68. Cited on page 46.

(Taylor, 1994) J. A. Taylor. 1994. *A multi-agent planner for modelling dialogue*. Ph.D. Thesis, School of Cognitive and Computing Sciences, University of Sussex. Cited on page 46.

(Traum and Allen, 1994) David Traum and James Allen. 1994. Discourse obligations in dialogue processing. In *Proceedings of the 32nd annual meeting of ACL*, pages 1–8. Association for Computational Linguistics Morristown, NJ, USA. Cited on pages 8, 33, 35, 36, 37, 41, 44, 48, 61, 74, 268, 271, and 277.

(Traum and Hinkelman, 1992) David Traum and Elizabeth A Hinkelman. 1992. Conversation acts in task-oriented spoken dialogue. *Computational intelligence*, 8(3):575–599. Cited on pages 33, 38, 69, 110, and 267.

(Traum and Larsson, 2003) David Traum and Staffan Larsson. 2003. The information state approach to dialogue management. *Current and New Directions in Discourse and Dialogue*, pages 325–353. Cited on pages 40, 219, and 267.

(Traum et al., 1999) David Traum, Johan Bos, Robin Cooper, Staffan Larsson, Ian Lewin, Colin Matheson, and Massimo Poesio. 1999. A model of dialogue moves and information state revision. Technical report, The TRINDI Project. Cited on page 267.

(Traum et al., 2005) David Traum, W. Swartout, S. Marsella, and J. Gratch. 2005. Fight, Flight, or Negotiate: Believable Strategies for Conversing Under Crisis. *Lecture Notes in Computer Science*, 3661:52. Cited on pages 46 and 72.

(Traum et al., 2007) David Traum, A. Roque, A. Leuski, P. Georgiou, J. Gerten, B. Martinovski, S. Narayanan, S. Robinson, and A. Vaswani. 2007. Hassan: A virtual human for tactical questioning. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, Belgium*, pages 71–74. Cited on page 47.

(Traum et al., 2008a) David Traum, W. Swartout, J. Gratch, and S. Marsella. 2008a. A virtual human dialogue model for non-team interaction. *Recent Trends in Discourse and Dialogue. Springer.* Cited on page 46.

(Traum et al., 2008b) David Traum, W. Swartout, J. Gratch, and S. Marsella. 2008b. A Virtual Human Dialogue Model for Non-team Interaction. In Laila Dybkjær and Wolfgang Minker, editors, *Recent Trends in Discourse and Dialogue*. Springer. Cited on page 72.

(Traum et al., 2012) David Traum, David DeVault, Jina Lee, Zhiyang Wang, and Stacy Marsella. 2012. Incremental dialogue understanding and feedback for multiparty, multimodal conversation. In *Intelligent Virtual Agents*, pages 275–288. Springer. Cited on page 71.

(Traum, 1994) David Traum. 1994. *A Computational Theory of Grounding in Natural Language Conversation.* Ph.D. thesis, Ph.D. thesis, Computer Science, University of Rochester, New York, December. Cited on pages 34, 71, and 73.

(Traum, 2003) David Traum. 2003. Semantics and Pragmatics of Questions and Answers for Dialogue Agents. In *Proceedings of the International Workshop on Computational Semantics (IWCS)*, pages 380–394, January. Cited on page 218.

(Traum, 2008) David Traum. 2008. Extended Abstract: Computational Models of Non-cooperative dialogue. In Jonathan Ginzburg, Patrick Healey, and Yo Sato, editors, *Proceedings of LONDIAL 2008, the 12th Workshop on the Semantics and Pragmatics of Dialogue*, pages 11–14, London, UK. Cited on pages 46 and 72.

(Traum, 2012) David Traum. 2012. Non-cooperative and Deceptive Virtual Agents. *IEEE Intelligent Systems: Trends and Controversies: Computational Deception and Noncooperation*, 27(6):66–69. Cited on page 47.

(Wahlster and Kobsa, 1986) W. Wahlster and A. Kobsa. 1986. Dialogue-based user models. *Proceedings of the IEEE*, 74(7):948–960. Cited on page 45.

(Wahlster, 1993) W. Wahlster. 1993. *VERBMOBIL-translation of face-to-face dialogs*. Springer-Verlag London, UK. Cited on page 32.

(Walton and Krabbe, 1995) Douglas Walton and E. Krabbe. 1995. *Commitment in dialogue: Basic concepts of interpersonal reasoning*. State University of New York Press. Cited on pages 6, 8, 27, 28, 43, 61, 62, 72, 73, and 74.

(Zinn et al., 2002) Claus Zinn, Johanna Moore, and Mark Core. 2002. A 3-Tier Planning Architecture for Managing Tutorial Dialogue. In *Intelligent Tutoring Systems*, pages 574–584. Springer Berlin Heidelberg. Cited on page 220.

Appendix A

Annotation Study Materials

This appendix contains the materials involved in the annotation study described in Section 4.3. These include the transcript of one of the interviews fragments in the corpus as given to the annotators in each stage, the annotation guidelines and a description of the annotation tool used in the study. Further materials are available at http://mcs.open.ac.uk/nlg/non-cooperation/.

A.1 Corpus of Political Interviews

The following six interview fragments were used in the annotation study:

Interview	**Turns**	**Words**
1. Brodie and Blair	16	734
2. Green and Miliband	9	526
3. O'Reilly and Hartman	19	360
4. Paxman and Osborne	16	272
5. Pym and Osborne	10	595
6. Shaw and Thatcher	18	1069
Total	88	3556

The transcripts were selected from a larger set of 15 interviews collected from publicly available sources (BBC News, CNN, Youtube, etc.)[1]. When

[1]Copyright of all media and transcripts belongs to the respective broadcasting company. Interviews 1, 2, 4 and 5 are property of the British Broadcasting Company (BBC). Interview 3 is property of Fox News Network, L.L.C. Interview 6 is property of Cable News Network, Inc. (CNN).

available, official transcripts from the original source were used, with minor modifications to reduce the number of functionally empty or split turns (e.g. due to interruptions or overlapped speech). Otherwise, the interviews were transcribed from video or audio taken from the source. The following table lists the sources for the interview fragments in the corpus[2]:

Interview	Source
1. Brodie and Blair	http://news.bbc.co.uk/1/hi/uk/1552265.stm
2. Green and Miliband	http://www.bbc.co.uk/news/uk-politics-13971770
3. O'Reilly and Hartman	http://mediamatters.org/items/200801220012
4. Paxman and Osborne	http://www.youtube.com/watch?v=XGWcSkCu69c
5. Pym and Osborne	http://www.bbc.co.uk/news/uk-12275973
6. Shaw and Thatcher	http://edition.cnn.com/WORLD/9706/30/thatcher.transcript/

The transcript and a brief description of the context of Interview 1 are given below as received by the annotators in each annotation stage. The rest of the corpus is available online at http://mcs.open.ac.uk/nlg/non-cooperation/.

A.1.1 First Annotation Stage: Segmenting Turns

In the first stage, annotators received the transcripts without any annotations, other than the division of turns as spoken by each speaker.

Interview 1: Brodie and Blair

Context. *Shortly after 11 September 2001, UK Prime Minister Tony Blair is interviewed by Alex Brodie for BBC World Service's Newshour on the role of the UK after the terrorist attacks.*

Transcript.

Turn	Spkr.	Speech
0	IR	Is Osama Bin Laden your prime suspect?
1	IE	He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.
2	IR	Him alone or anybody else?
3	IE	Well, when we assemble the evidence finally, we will present it to people. But as we have said he is the prime suspect.
4	IR	Have you seen evidence yourself?

[2]Online sources were last accessed in August 2013.

Turn	Spkr.	Speech
5	IE	Yes of course, all the time we are going through evidence that comes to us from various sources and what is important, as I said the other day, is that when we proceed, we proceed on the basis of a hard-headed assessment of that evidence. But I think, people are still taking in the enormity of what happened last week. Thousands of people killed in the worst terrorist incident of all time. This was the worst terrorist incident in respect of British citizens, incidentally 200, 300 killed, since World War II. When you think that Britain went through the Blitz when we were under attack, day in day out, for several years and we lost just over 20,000 of our citizens. Here were 5,000 or more murdered, literally, in a day and I think some impression is given of just how serious this is. Let's be quite clear as well, the thing that we have to confront and the reason why we have to take action against this apparatus of terrorism at every level, is that if these people were able to kill more people they would. The only limits on their actions are not moral in any sense at all, they are practical or technical.
6	IR	Is it Osama Bin Laden who you have the evidence against that he was actively involved in planning what happened in the United States or is it just that you have evidence that he has set up a network?
7	IE	Well Alex, when we are in a position to put evidence before people, we will put it before them then. What we have said so far, because people have asked us and it's right because this is where the evidence tends, that he is the prime suspect.
8	IR	Anybody else?
9	IE	There may be various other people but that is a matter that we can deal with when we come to present the evidence fully.
10	IR	And do you know where he is?
11	IE	We know that he is in Afghanistan. We know the various places that he has been. But it is important that other people co-operate with us in ensuring that he is brought to justice and this is a situation in which those who have been harbouring him or helping him have a very simple choice. They either cease the protection of Bin Laden or they will be treated as people helping him.
12	IR	This is echoing what George Bush said isn't it about how we will go not just for the perpetrators but for those who harbour him, and you are talking about the Taliban?
13	IE	Well, for all those people who have been in a position where they have been helping or harbouring terrorism, the way that it operates, camps that are dedicated to training people in it. These are people trained in these camps who go out and basically wreak havoc wherever they can, killing many, many innocent people. And although what happened last week is obviously an atrocity almost beyond our imagination, it is not an isolated incident, in that sense, there has been a history going back over several years. Now you mention the Taliban, the Taliban have a very clear choice, the Taliban either cease to help or harbour those that are fermenting terrorism or they will be treated as part of the terrorist apparatus themselves. Now they have that choice and they should consider very, very carefully the consequences that they face at this moment of choice.
14	IR	If they don't give him up, what are those consequences?
15	IE	Those are the consequences again that we will consider and we will announce the appropriate response when we have made up our minds.

A.1.2 Second Annotation Stage: Selecting Content Features

In the second stage, annotators received the dialogues segmented and annotated with dialogue act functions and, when applicable, referent segments. These partial annotations were obtained automatically as described in Section 4.3.3.

Legend. Partial annotations are marked directly on the speech transcript. Segment boundaries are indicated using square brackets and numbered sequentially. Inside the opening brackets, dialogue act functions are shown according to the following key:

1 **Init-Inform**
2 **Init-InfoReq**
3 **Resp-Inform**
4 **Resp-Accept**
5 **Resp-Reject**

Referent segments are indicated after the dialogue act function using the "@" symbol and the number of the segment they point to. The following marking in Interview 1, for example, identifies segment **(4)**, with dialogue act function **Resp-Inform** and referent segment **(2)**:

(4)[**3@(2)** But as we have said he is the prime suspect.]

Interview 1: Brodie and Blair

Context. *Shortly after 11 September 2001, UK Prime Minister Tony Blair is interviewed by Alex Brodie for BBC World Service's Newshour on the role of the UK after the terrorist attacks.*

Segmented and Partially Annotated Transcript.

Turn	Spkr.	Annotated Speech
0	IR	**(0)**[**2** Is Osama Bin Laden your prime suspect?]
1	IE	**(1)**[**3@(0)** He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.]
2	IR	**(2)**[**2** Him alone or anybody else?]
3	IE	**(3)**[**3@(2)** Well, when we assemble the evidence finally, we will present it to people.] **(4)**[**3@(2)** But as we have said he is the prime suspect.]
4	IR	**(5)**[**2** Have you seen evidence yourself?]

Turn	Spkr.	Annotated Speech
5	IE	**(6)[3@(5)** Yes of course, all the time we are going through evidence that comes to us from various sources and what is important, as I said the other day, is that when we proceed, we proceed on the basis of a hard-headed assessment of that evidence.] **(7)[3@(5)** But I think, people are still taking in the enormity of what happened last week. Thousands of people killed in the worst terrorist incident of all time. This was the worst terrorist incident in respect of British citizens, incidentally 200, 300 killed, since World War II. When you think that Britain went through the Blitz when we were under attack, day in day out, for several years and we lost just over 20,000 of our citizens. Here were 5,000 or more murdered, literally, in a day and I think some impression is given of just how serious this is.] **(8)[3@(5)** Let's be quite clear as well, the thing that we have to confront and the reason why we have to take action against this apparatus of terrorism at every level, is that if these people were able to kill more people they would. The only limits on their actions are not moral in any sense at all, they are practical or technical.]
6	IR	**(9)[2** Is it Osama Bin Laden who you have the evidence against that he was actively involved in planning what happened in the United States or is it just that you have evidence that he has set up a network?]
7	IE	**(10)[3@(9)** Well Alex, when we are in a position to put evidence before people, we will put it before them then. What we have said so far, because people have asked us and it's right because this is where the evidence tends, that he is the prime suspect.]
8	IR	**(11)[2** Anybody else?]
9	IE	**(12)[3@(1)1** There may be various other people but that is a matter that we can deal with when we come to present the evidence fully.]
10	IR	**(13)[2** And do you know where he is?]
11	IE	**(14)[3@(1)3** We know that he is in Afghanistan. We know the various places that he has been.] **(15)[3@(1)3** But it is important that other people co-operate with us in ensuring that he is brought to justice and this is a situation in which those who have been harbouring him or helping him have a very simple choice. They either cease the protection of Bin Laden or they will be treated as people helping him.]
12	IR	**(16)[1** This is echoing what George Bush said isn't it about how we will go not just for the perpetrators but for those who harbour him] , **(17)[2** and you are talking about the Taliban?]
13	IE	**(18)[3@(1)7** Well, for all those people who have been in a position where they have been helping or harbouring terrorism, the way that it operates, camps that are dedicated to training people in it. These are people trained in these camps who go out and basically wreak havoc wherever they can, killing many, many innocent people. And although what happened last week is obviously an atrocity almost beyond our imagination, it is not an isolated incident, in that sense, there has been a history going back over several years.] **(19)[3@(1)7** Now you mention the Taliban, the Taliban have a very clear choice, the Taliban either cease to help or harbour those that are fermenting terrorism or they will be treated as part of the terrorist apparatus themselves. Now they have that choice and they should consider very, very carefully the consequences that they face at this moment of choice.]
14	IR	**(20)[2** If they don't give him up, what are those consequences?]
15	IE	**(21)[3@(2)0** Those are the consequences again that we will consider and we will announce the appropriate response when we have made up our minds.]

A.2 Annotation Tool

The annotation was carried out using a special-purpose tool[3], deployed to each annotator containing the annotation data. Among other features, the tool guides the annotators through the dataset in a fixed order and can be configured to operate according to each annotation stage.

The main window of the tool (Figure A.1) shows the interview context, the turn transcripts and the annotations. Clicking on the annotation next to a turn opens a window that allows the to segment and annotate the turn. The tool, user guide and annotated examples are available at http://mcs.open.ac.uk/nlg/non-cooperation/.

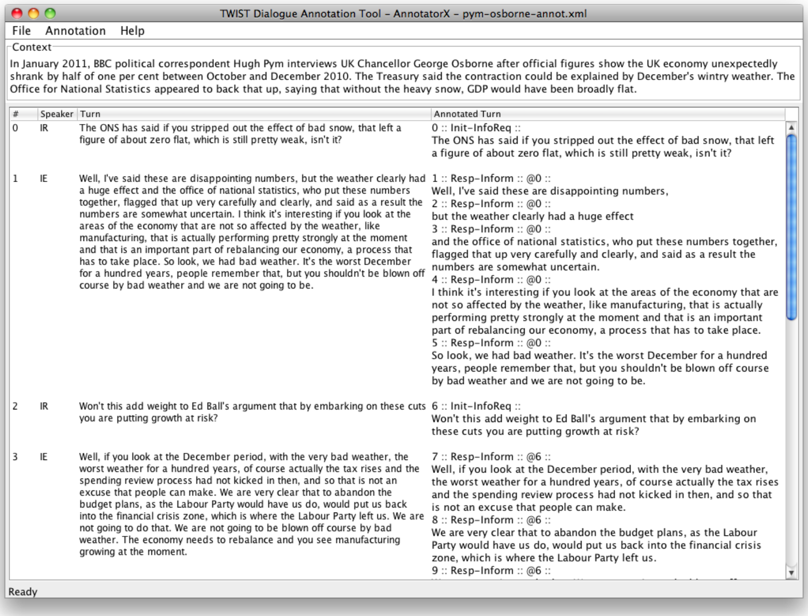


Figure A.1: Annotation Tool (main window)

A.3 Annotation Guidelines

Annotators were given a set of guidelines for each stage with the definitions and examples presented in Section 4.2. These can can be found at http://mcs.open.ac.uk/nlg/non-cooperation/. Together with the guidelines, annotators received a brief introduction to the instructions and examples before starting their annotations. They were asked to read the document in detail and had a chance to ask questions about anything that needed clarification.

[3]The tool was built based on the CODA *D2MTool* developed by Svetlana Stoyanchev for the CODA Project (Stoyanchev and Piwek, 2010b).

Appendix B

Measuring Cooperation in Annotated Dialogue: Output for Interview 1

The method for automatically measuring the degree of cooperation and non-cooperation of each speaker in annotated political interviews described in Chapter 4 has been implemented in Java and applied to a corpus of annotated interviews for evaluation. This appendix presents the output produced by the implementation. It corresponds to Interview 1 in the corpus, which is used as an example throughout Section 4.4.2 for illustration of the method:

```
(0) [(IR, valid-question)]

(1)
IR (0) valid-question
** Obligation (IR, valid-question) explicitly discharged by action (0) valid-question

OBL (1) [(IE, acceptance@0)]
DCF (1) [(IR, valid-question)]
DNF (1) []
SCF (1) [(0) valid-question ;]
SNF (1) []

(2)
IE (1) valid-reply @0 (C)
** Acknowledgment obligation (IE, acceptance@0) implicitly discharged by action (1) valid-reply @0 (C)
** Implicit obligation (IE, valid-reply@0) introduced.
** Obligation (IE, valid-reply@0) explicitly discharged by action (1) valid-reply @0 (C)

OBL (2) [(IR, acceptance@1)]
DCF (2) [(IE, acceptance@0),(IE, valid-reply@0)]
DNF (2) []
SCF (2) [(1) valid-reply @0 (C) ;]
SNF (2) []

(3)
IR (2) valid-question
** Acknowledgment obligation (IR, acceptance@1) implicitly discharged by action (2) valid-question
** Implicit obligation (IR, valid-question) introduced.
** Obligation (IR, valid-question) explicitly discharged by action (2) valid-question
```

```
OBL (3) [(IE, acceptance@2)]
DCF (3) [(IR, acceptance@1),(IR, valid-question)]
DNF (3) []
SCF (3) [(2) valid-question ;]
SNF (3) []

(4)
IE (3) valid-reply @2 (I)
IE (4) invalid-reply @2 (R) (I) {Reason: Repeated}
** Acknowledgment obligation (IE, acceptance@2) implicitly discharged by action (3) valid-reply @2 (I)
** Implicit obligation (IE, valid-reply@2) introduced.

OBL (4) [(IR, acceptance@3),(IR, rejection@4),(IE, valid-reply@2)]
DCF (4) [(IE, acceptance@2)]
DNF (4) [(IE, valid-reply@2)]
SCF (4) [(3) valid-reply @2 (I) ;]
SNF (4) [(4) invalid-reply @2 (R) (I) {Reason: Repeated};]

(5)
IR (5) valid-question
** Acknowledgment obligation (IR, acceptance@3) implicitly discharged by action (5) valid-question
** Implicit obligation (IR, valid-question) introduced.
** Obligation (IR, valid-question) explicitly discharged by action (5) valid-question

OBL (5) [(IE, acceptance@5),(IR, rejection@4),(IE, valid-reply@2)]
DCF (5) [(IR, acceptance@3),(IR, valid-question)]
DNF (5) [(IR, rejection@4)]
SCF (5) [(5) valid-question ;]
SNF (5) []

(6)
IE (6) valid-reply @5 (C)
IE (7) invalid-reply @5 (C) {Reason: Irrelevant}
IE (8) invalid-reply @5 (C) {Reason: Irrelevant}
** Acknowledgment obligation (IE, acceptance@5) implicitly discharged by action (6) valid-reply @5 (C)
** Implicit obligation (IE, valid-reply@5) introduced.
** Obligation (IE, valid-reply@5) explicitly discharged by action (6) valid-reply @5 (C)

OBL (6) [(IR, acceptance@6),(IR, rejection@7),(IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (6) [(IE, acceptance@5),(IE, valid-reply@5)]
DNF (6) [(IE, valid-reply@2)]
SCF (6) [(6) valid-reply @5 (C) ;]
SNF (6) [(7) invalid-reply @5 (C) {Reason: Irrelevant}; (8) invalid-reply @5 (C) {Reason: Irrelevant};]

(7)
IR (9) valid-question
** Acknowledgment obligation (IR, acceptance@6) implicitly discharged by action (9) valid-question
** Implicit obligation (IR, valid-question) introduced.
** Obligation (IR, valid-question) explicitly discharged by action (9) valid-question

OBL (7) [(IE, acceptance@9),(IR, rejection@7),(IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (7) [(IR, acceptance@6),(IR, valid-question)]
DNF (7) [(IR, rejection@7),(IR, rejection@8),(IR, rejection@4)]
SCF (7) [(9) valid-question ;]
SNF (7) []

(8)
IE (10) invalid-reply @9 (R) (I) {Reason: Repeated}
** Acknowledgment obligation (IE, acceptance@9) implicitly discharged
   by action (10) invalid-reply @9 (R) (I) {Reason: Repeated}
** Implicit obligation (IE, valid-reply@9) introduced.

OBL (8) [(IR, rejection@10),(IE, valid-reply@9),(IR, rejection@7),(IR, rejection@8),
         (IR, rejection@4),(IE, valid-reply@2)]
DCF (8) [(IE, acceptance@9)]
DNF (8) [(IE, valid-reply@9),(IE, valid-reply@2)]
SCF (8) []
SNF (8) [(10) invalid-reply @9 (R) (I) {Reason: Repeated};]

(9)
IR (11) repeated-valid-question (R)
** Acknowledgment obligation (IR, rejection@10) implicitly discharged
   by action (11) repeated-valid-question (R)

OBL (9) [(IE, valid-reply@9),(IR, rejection@7),(IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (9) [(IR, rejection@10)]
DNF (9) [(IR, rejection@7),(IR, rejection@8),(IR, rejection@4)]
SCF (9) [(11) repeated-valid-question (R) ;]
SNF (9) []

(10)
IE (12) invalid-reply @11 (R) (I) {Reason: Repeated}

OBL (10) [(IR, rejection@12),(IE, valid-reply@9),(IR, rejection@7),(IR, rejection@8),
```

```
             (IR, rejection@4),(IE, valid-reply@2)]
DCF (10) []
DNF (10) [(IE, valid-reply@9),(IE, valid-reply@2)]
SCF (10) []
SNF (10) [(12) invalid-reply @11 (R) (I) {Reason: Repeated};]

(11)
IR (13) valid-question
** Implicit obligation (IR, valid-question) introduced.
** Obligation (IR, valid-question) explicitly discharged by action (13) valid-question

OBL (11) [(IE, acceptance@13),(IR, rejection@12),(IE, valid-reply@9),(IR, rejection@7),(IR, rejection@8),
             (IR, rejection@4),(IE, valid-reply@2)]
DCF (11) [(IR, valid-question)]
DNF (11) [(IR, rejection@12),(IR, rejection@7),(IR, rejection@8),(IR, rejection@4)]
SCF (11) [(13) valid-question ;]
SNF (11) []

(12)
IE (14) valid-reply @13 (C)
IE (15) invalid-reply @13 (C) {Reason: Irrelevant}
** Acknowledgment obligation (IE, acceptance@13) implicitly discharged by action (14) valid-reply @13 (C)
** Implicit obligation (IE, valid-reply@13) introduced.
** Obligation (IE, valid-reply@13) explicitly discharged by action (14) valid-reply @13 (C)

OBL (12) [(IR, acceptance@14),(IR, rejection@15),(IR, rejection@12),(IE, valid-reply@9),(IR, rejection@7),
             (IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (12) [(IE, acceptance@13),(IE, valid-reply@13)]
DNF (12) [(IE, valid-reply@9),(IE, valid-reply@2)]
SCF (12) [(14) valid-reply @13 (C) ;]
SNF (12) [(15) invalid-reply @13 (C) {Reason: Irrelevant};]

(13)
IR (16) valid-statement
IR (17) valid-question
** Acknowledgment obligation (IR, acceptance@14) implicitly discharged by action (16) valid-statement
** Implicit obligation (IR, valid-question) introduced.
** Obligation (IR, valid-question) explicitly discharged by action (17) valid-question

OBL (13) [(IE, acceptance@16),(IE, acceptance@17),(IR, rejection@15),(IR, rejection@12),(IE, valid-reply@9),
             (IR, rejection@7),(IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (13) [(IR, acceptance@14),(IR, valid-question)]
DNF (13) [(IR, rejection@15),(IR, rejection@12),(IR, rejection@7),(IR, rejection@8),(IR, rejection@4)]
SCF (13) [(16) valid-statement ; (17) valid-question ;]
SNF (13) []

(14)
IE (18) valid-reply @17 (I)
IE (19) valid-reply @17 (C)
** Acknowledgment obligation (IE, acceptance@16) implicitly discharged by action (18) valid-reply @17 (I)
** Acknowledgment obligation (IE, acceptance@17) implicitly discharged by action (18) valid-reply @17 (I)
** Implicit obligation (IE, valid-reply@17) introduced.
** Obligation (IE, valid-reply@17) explicitly discharged by action (19) valid-reply @17 (C)

OBL (14) [(IR, acceptance@18),(IR, acceptance@19),(IR, rejection@15),(IR, rejection@12),(IE, valid-reply@9),
             (IR, rejection@7),(IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (14) [(IE, acceptance@16),(IE, acceptance@17),(IE, valid-reply@17)]
DNF (14) [(IE, valid-reply@9),(IE, valid-reply@2)]
SCF (14) [(18) valid-reply @17 (I) ; (19) valid-reply @17 (C) ;]
SNF (14) []

(15)
IR (20) valid-question
** Acknowledgment obligation (IR, acceptance@18) implicitly discharged by action (20) valid-question
** Acknowledgment obligation (IR, acceptance@19) implicitly discharged by action (20) valid-question
** Implicit obligation (IR, valid-question) introduced.
** Obligation (IR, valid-question) explicitly discharged by action (20) valid-question

OBL (15) [(IE, acceptance@20),(IR, rejection@15),(IR, rejection@12),(IE, valid-reply@9),(IR, rejection@7),
             (IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (15) [(IR, acceptance@18),(IR, acceptance@19),(IR, valid-question)]
DNF (15) [(IR, rejection@15),(IR, rejection@12),(IR, rejection@7),(IR, rejection@8),(IR, rejection@4)]
SCF (15) [(20) valid-question ;]
SNF (15) []

(16)
IE (21) valid-reply @20 (I)
** Acknowledgment obligation (IE, acceptance@20) implicitly discharged by action (21) valid-reply @20 (I)
** Implicit obligation (IE, valid-reply@20) introduced.
New obligations not introduced (last turn): [(IR, acceptance@21)]

OBL (16) [(IE, valid-reply@20),(IR, rejection@15),(IR, rejection@12),(IE, valid-reply@9),(IR, rejection@7),
             (IR, rejection@8),(IR, rejection@4),(IE, valid-reply@2)]
DCF (16) [(IE, acceptance@20)]
```

```
DNF (16) [(IE, valid-reply@20),(IE, valid-reply@9),(IE, valid-reply@2)]
SCF (16) [(21) valid-reply @20 (I) ;]
SNF (16) []

DC for IR: 0.5227272727272727 (DCFs: 14.0; DNFs:21.0; SCFs: 9.0; SNFs:0.0)
DC for IE: 0.5 (DCFs: 12.0; DNFs:13.0; SCFs: 7.0; SNFs:6.0)
```

Appendix C

Survey for Eliciting Human Judgement of Cooperation in Dialogue

This appendix presents the online survey described in Chapter 4 for eliciting human judgement on cooperation in political interviews.

C.1 Facsimile of Online Survey

Below is a reproduction of the survey as it was presented to participants on the SurveyMonkey site[1]:

[1] http://www.surveymonkey.com/

1. Introduction Exit this survey

This is a survey on how people perceive verbal behaviour in political interviews.

On each page you will be shown some context and a transcribed interview fragment and then asked to rate the performance of each participant according to your intuitions on how they ought to behave in a political interview.

Although you might recognise the characters in the interviews, please be as objective as possible when rating them, regardless of your personal feelings about their style or political views.

The entire survey should take about 20 minutes to complete, but if you do not have that amount of time to spare, please complete as much of it as you can and then skip through the rest until you get to the final page.

At the end of the survey, you can watch a few rather amusing interactions I have come across during my research.

Thanks for your help!

Next

2. Information Exit this survey

Before you start the survey, please answer these questions about your background.

***Which of the following best describes your English proficiency?**

- Native English speaker
- Non-native English speaker but fluent in English
- Non-native English speaker but know some English

***Which of the following best describes your cultural background?**

- British or American
- Neither British nor American, but have lived in the UK/US for many years (more than 5)
- Neither British nor American, but have lived in the UK/US for a few years (between 1 and 5)
- Neither British nor American, but Commonwealth
- Other

***Which of the following best describes your experience in dialogue analysis?**

- Expert
- Some experience in research
- Some informal experience
- No experience whatsoever

Next

3. Bernard Shaw and Margaret Thatcher Exit this survey

Context

On Sunday 29 June 1997, CNN News anchor Bernard Shaw interviews former British Prime Minister Margaret Thatcher in the context of the transfer of sovereignty over Hong Kong from the United Kingdom to China.

Shaw What is the difference between negotiation, say, with the Russians and the Chinese?

Thatcher Well, right now, Russia proved what we always said would happen, although it came quicker than we thought. We knew the communist system eventually would collapse. You can't ignore human rights eventually, without the system collapsing, particularly in the modern world where they can't keep out information on the Internet about what's happening to other countries. And also, Mr. Gorbachev, he doesn't get enough credit, realised the communist system wasn't working economically, was not producing prosperity, was meant to be the system that produced the greatest prosperity because it was all planned. It doesn't produce prosperity because it offers no stimulus or incentive to people to build up their own prosperity. So it came faster in Russia. China has no history of liberty at all. She has always been under tyranny. She went from being under Chiang Kai Shek and Kuomintang, to come under communism in 1949. It will eventually collapse also.

Shaw Do you think this system of government here in China-

Thatcher (Interrupting) Communsim will eventually collapse. Indeed, it is starting. Deng Xiaoping realized it couldn't go on. So he said right, economic liberty. You can start up your own business. If you produce more than your target in the factories you can set out to sell it. They are born traders the Chinese. Beijing is so different from what it was in 1977. It has got the economic liberty. It has not yet got a full rule of law, although they are having to supply now and create a law of contract so that you can in fact enforce your own contract. Law is coming too, to China, initiative is coming to China, enterprise is coming to China. It won't stop.

Shaw Might things have been better had there been better chemistry between you and Deng Xiaoping? During the 1982 talks, referring to you, Mr. Deng said 'that woman should be bombarded out of her obstinance.'

Thatcher Well, that is what he'd want to say, wouldn't he? If you had argued with him you are obstinate. He was obstinate, he argued with me. But I didn't complain about that. We survive on argument, that is how come to the right conclusions. Yes, I was obstinate and because of that at any rate we didn't get a good agreement because of dependent detail. Because he knew we produced prosperity and he didn't and he started to change. Why? Of course, I am obstinate in defending our liberties and our law. That is why I carry a big handbag.

Shaw Following the Falklands War, did hubris from having won that war make you believe that you could persuade the Chinese that Britain should continue administering Hong Kong with an umbrella of Chinese sovereignty?

Thatcher No, there was no hubris in Falklands, only a fantastic relief that our people were once again free and we were not going to have an aggressor taking over British land and British people. And we don't like aggression anywhere in the world, that is why we believe in strong defense.

Shaw Well, Sir Percy Craddock, Britain's Ambassador to China said that you had to be persuaded, that you had to be told, that there was no way Britain was going to remain an administrative force of Hong Kong with the Chinese being the mere sovereigns.

Thatcher Well, that Deng Xiaoping told me. I'll tell you what he told me. I have written it. I said that we have done so well for Hong Kong, for Hong Kong people, that can we not have another lease say for another 50 years? He reacted very quickly. He said no. I said can we not have another lease? I said we have done so well on a territory which I know will eventually return to you. Wouldn't you really let us have, it would be an act of sovereignty to give us a management contract?

Shaw They were outraged. Is that when Mr. Deng told you that if the Chinese wanted to they could walk right in here and take Hong Kong?

Thatcher Oh yes he said he could. But I know that I didn't need to be told. That is why I had to ask him. But, he said to me, which really rather shook me 'I would rather recover Hong Kong poverty stricken than let the British have another period of administration over Hong Kong.' Now, that shows you the communist mind, not concerned about the prosperity, about the well being of the people.

Shaw You don't trust him, do you?

Thatcher I don't trust a communist, do you?

Shaw I can't answer that, I am the reporter asking questions.

Thatcher It is interesting that you asked it. Just make an assessment of the person you are negotiating with. What I had to do was, I knew that Hong Kong was valuable to him. I knew that they could do a lot through Hong Kong that they couldn't do otherwise. And so eventually he agreed. And when he said to me 'I could take it over, I could take it over this afternoon', I said 'yes, you could, and it would become poverty stricken, because there would be alarm, people would leave, and the world would know it was the dead hand of communism that ruined it.' So, he said 'what did you have on that piece of paper, Mrs. Thatcher?' And I had written out a possible communique which said that we had decided to negotiate about the future of Hong Kong. Perhaps not that we'd negotiate that we'd have a series of meetings about the matters that would come up. This is 15 years, because we could not get any loans from banks for properties, anyone, in less than 15 years, so we had to negotiate. And we did the communique which I had drafted and the negotiations started and it took two years.

Shaw At these historic ceremonies, will you be fighting back tears?

Thatcher I hope the tears won't flow. My mind and heart will just be very full for the people of Hong Kong. And just tremendous hope that all will be well, and a determination that, along with other democratic countries in the world, we observe very carefully what is going on in Hong Kong. And we don't hesitate to speak out for the people of Hong Kong and do what we can to see that that international agreement I made with Deng Xiaoping, registered in the United Nations, is fully observed and upheld.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in <u>this</u> fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
Shaw	○	○	○	○	○
Thatcher	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

Next

4. Damon Green and Ed Miliband Exit this survey

Context

In June 2011, ITV News correspondent Damon Green interviewed UK Labour leader Ed Miliband on his position regarding a strike action organised by public sector workers. The action was a protest against planned pension changes. The strike action resulted in the closure of almost half of the state schools across the UK. The interview starts with Miliband stating his position with regards the matter.

Miliband These strikes are wrong at a time when negotiations are still going on. But parents and the public have been let down by both sides because the government has acted in a reckless and provocative manner. After today's disruption, I urge both sides to put aside the rhetoric, get round the negotiating table and stop it happening again.

Green I listened to your speech in Wrexham where you talked about the Labour Party being a movement. A lot of people in that movement are the people who are on strike today and they'll be looking at you and thinking 'well, you're describing these strikes as wrong, why aren't you giving us more leadership as the leader of the Labour movement?'

Miliband At a time when negotiations are still going on I do believe these strikes are wrong. And that's why I say both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and sort the problem out. Because the public and parents have been let down by both sides. The government has acted in a reckless and provocative manner.

Green I spoke to Francis Maude before I came here and the tone he was striking was a very conciliatory one. Do you think there's a difference between the words they are saying in public and the attitudes they strike in private behind the negotiations? Are the negotiations in good faith would you say?

Miliband What I say is that the strikes are wrong at a time when negotiations are still going on. But the government has acted in a reckless and provocative manner in the way it has gone about these issues. After today's disruption, I urge both sides to get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.

Green It's a- it's a statement you've made publicly, and you've made to me and this will be broadcast, obviously, but have you spoken privately to any union leaders and expressed your view to them on a personal level, would you say?

Miliband What I say in public and in private, to everybody involved in this, is get round the negotiating table, put aside the rhetoric, and stop this kind of action happening again. These strikes are wrong because negotiations are still going on. But parents and the public have been let down by the government as well, who've acted in a reckless and provocative manner.

Green You're a parent. I'm a parent. People who will be watching this are parents. Umm, has it affected you personally, this action? Has it affected your family, your friends, I mean? What is the net effect of that going to be on parents having to take a day off work today?

Miliband I think parents up and down the country have been affected by this action, and it's wrong at a time when negotiations are still going on. Parents have been let down by both sides because the government has acted in a reckless and provocative manner. I think that both sides should, after today's disruption, get round the negotiating table, put aside the rhetoric, and stop this kind of thing happening again.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in <u>this</u> fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
Green	○	○	○	○	○
Miliband	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

Next

5. Alex Brodie and Tony Blair

Exit this survey

Context

Shortly after 11 September 2001, UK Prime Minister Tony Blair is interviewed by Alex Brodie for BBC World Service's Newshour on the role of the UK after the terrorist attacks.

Brodie Is Osama Bin Laden your prime suspect?

Blair He is the prime suspect. We are still assembling the evidence and we have said we will do so in a careful and measured way. But we've known for some time of his activities and those of his associates, that have been designed to spread terror around the world that are I believe fundamentally contrary to the basic teachings of Islam. And in respect of this particular incident there's no doubt at all, as both ourselves and President have said, he is the prime suspect.

Brodie Him alone or anybody else?

Blair Well, when we assemble the evidence finally, we will present it to people. But as we have said he is the prime suspect.

Brodie Have you seen evidence yourself?

Blair Yes of course, all the time we are going through evidence that comes to us from various sources and what is important, as I said the other day, is that when we proceed, we proceed on the basis of a hard-headed assessment of that evidence. But I think, people are still taking in the enormity of what happened last week. Thousands of people killed in the worst terrorist incident of all time. This was the worst terrorist incident in respect of British citizens, incidentally 200, 300 killed, since World War II. When you think that Britain went through the Blitz when we were under attack, day in day out, for several years and we lost just over 20,000 of our citizens. Here were 5,000 or more murdered, literally, in a day and I think some impression is given of just how serious this is. Let's be quite clear as well, the thing that we have to confront and the reason why we have to take action against this apparatus of terrorism at every level, is that if these people were able to kill more people they would. The only limits on their actions are not moral in any sense at all, they are practical or technical.

Brodie Is it Osama Bin Laden who you have the evidence against that he was actively involved in planning what happened in the United States or is it just that you have evidence that he has set up a network?

Blair Well Alex, when we are in a position to put evidence before people, we will put it before them then. What we have said so far, because people have asked us and it's right because this is where the evidence tends, that he is the prime suspect.

Brodie Anybody else?

Blair There may be various other people but that is a matter that we can deal with when we come to present the evidence fully.

Brodie And do you know where he is?

Blair We know that he is in Afghanistan. We know the various places that he has been. But it is important that other people co-operate with us in ensuring that he is brought to justice and this is a situation in which those who have been harbouring him or helping him have a very simple choice. They either cease the protection of Bin Laden or they will be treated as people helping him.

Brodie This is echoing what George Bush said isn't it about how we will go not just for the perpetrators but for those who harbour him, and you are talking about the Taliban?

Blair Well, for all those people who have been in a position where they have been helping or harbouring terrorism, the way that it operates, camps that are dedicated to training people in it. These are people trained in these camps who go out and basically wreak havoc wherever they can, killing many, many innocent people. And although what happened last week is obviously an atrocity almost beyond our imagination, it is not an isolated incident, in that sense, there has been a history going back over several years. Now you mention the Taliban, the Taliban have a very clear choice, the Taliban either cease to help or harbour those that are fermenting terrorism or they will be treated as part of the terrorist apparatus themselves. Now they have that choice and they should consider very, very carefully the consequences that they face at this moment of choice.

Brodie If they don't give him up, what are those consequences?

Blair Those are the consequences again that we will consider and we will announce the appropriate response when we have made up our minds.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in <u>this</u> fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
Brodie	○	○	○	○	○
Blair	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

6. Bill O'Reilly and Hermene Hartman

Exit this survey

Context

During the American Presidential campaign in January 2008, Fox News host Bill O'Reilly interviews Hermene Hartman, the editor of an African-American newspaper in Chicago. The interview is about Obama's pastor Jeremiah Wright and his connections with Nation of Islam's leader Louis Farrakhan.

O'Reilly How would you describe Dr Wright's church?

Hartman It's a middle-class church. It is a superb church. Reverend Wright started a church with 87 people; today, has 8,000 in that particular congregation. United Church of Christ is basically a white denomination. And I think there's been just a lot of miscasting here. Seventy ministries within the church, to include Girl Scouts, prison outreach, marital counselling, education, children's counselling, a lot of Adopt-A-School. They have done a lot to empower that community and to improve that community.

O'Reilly OK. But you could make the same argument about Louis Farrakhan, that he's done, you know, some good things, yet you know, his anti-semitic in his rhetoric and sometimes anti-white or whatever. And-

Hartman (Interrupting) But that is, that is not Jeremiah Wright.

O'Reilly No, but it is association there. And the association, you can draw your own conclusion.

Hartman But what - what's the emphasis? I mean, you could also, you know, it's the twist. It's the turn that's being taken. You could also look at a wonderful sermon that Dr Wright gave and a book developed out of it, The Audacity of Hope.

O'Reilly But you can't, you can't do that, though.

Hartman But we're, but here's what, you can do that if you wanted to do that.

O'Reilly No, no, no, no.

Hartman (Overlapping) You could. Here's what, but Bill-

O'Reilly (Overlapping) Because every despot, and I'm not calling the man a despot, but every despot in history has done some good things. Here, look-

Hartman (Interrupting) But he's not a despot. Come on, Bill.

O'Reilly No, I'm not, I'm not calling him that.

Hartman That's, that's out of order.

O'Reilly I made that clear.

Hartman (Overlapping) Well, what are you saying?

O'Reilly (Overlapping) But the things that he has said are very, very troubling. And I think that Senator Obama, if he's going to continue to associate with the Doctor, and he says he will-

Hartman (Interrupting) Obama is a- is running against a political couple. That is what is going on now. And true enough, obviously he's got to be judged just like everybody else, but you've got to bring the truth. If you're going to do Obama's church, let's do everybody's church.

O'Reilly All right.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in this fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
O'Reilly	○	○	○	○	○
Hartman	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

Next

7. Bill O'Reilly and Hermene Hartman Exit this survey

Context

During the American Presidential campaign in January 2008, Fox News host Bill O'Reilly interviews Hermene Hartman, the editor of an African-American newspaper in Chicago. The interview is about Obama's pastor Jeremiah Wright and his connections with Nation of Islam's leader Louis Farrakhan.

O'Reilly How would you describe Dr Wright's church?

Hartman It's a middle-class church. It is a superb church. Reverend Wright started a church with 87 people; today, has 8,000 in that particular congregation. United Church of Christ is basically a white denomination. And I think there's been just a lot of miscasting here. Seventy ministries within the church, to include Girl Scouts, prison outreach, marital counselling, education, children's counselling, a lot of Adopt-A-School. They have done a lot to empower that community and to improve that community.

O'Reilly OK. But you could make the same argument about Louis Farrakhan, that he's done, you know, some good things, yet you know, his anti-semitic in his rhetoric and sometimes anti-white or whatever. And-

Hartman (Interrupting) But that is, that is not Jeremiah Wright.

O'Reilly No, but it is association there. And the association, you can draw your own conclusion.

Hartman But what - what's the emphasis? I mean, you could also, you know, it's the twist. It's the turn that's being taken. You could also look at a wonderful sermon that Dr Wright gave and a book developed out of it, The Audacity of Hope.

O'Reilly But you can't, you can't do that, though.

Hartman But we're, but here's what, you can do that if you wanted to do that.

O'Reilly No, no, no, no.

Hartman (Overlapping) You could. Here's what, but Bill-

O'Reilly (Overlapping) Because every despot, and I'm not calling the man a despot, but every despot in history has done some good things. Here, look-

Hartman (Interrupting) But he's not a despot. Come on, Bill.

O'Reilly No, I'm not, I'm not calling him that.

Hartman That's, that's out of order.

O'Reilly I made that clear.

Hartman (Overlapping) Well, what are you saying?

O'Reilly (Overlapping) But the things that he has said are very, very troubling. And I think that Senator Obama, if he's going to continue to associate with the Doctor, and he says he will-

Hartman (Interrupting) Obama is a- is running against a political couple. That is what is going on now. And true enough, obviously he's got to be judged just like everybody else, but you've got to bring the truth. If you're going to do Obama's church, let's do everybody's church.

O'Reilly All right.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in this fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
O'Reilly	○	○	○	○	○
Hartman	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

Next

8. Hugh Pym and George Osborne

Exit this survey

Context

In January 2011, BBC political correspondent Hugh Pym interviews UK Chancellor George Osborne after official figures show the UK economy unexpectedly shrank by half of one per cent between October and December 2010. The Treasury said the contraction could be explained by December's wintry weather. The Office for National Statistics appeared to back that up, saying that without the heavy snow, GDP would have been broadly flat.

Pym The ONS has said if you stripped out the effect of bad snow, that left a figure of about zero flat, which is still pretty weak, isn't it?

Osborne Well, I've said these are disappointing numbers, but the weather clearly had a huge effect and the office of national statistics, who put these numbers together, flagged that up very carefully and clearly, and said as a result the numbers are somewhat uncertain. I think it's interesting if you look at the areas of the economy that are not so affected by the weather, like manufacturing, that is actually performing pretty strongly at the moment and that is an important part of rebalancing our economy, a process that has to take place. So look, we had bad weather. It's the worst December for a hundred years, people remember that, but you shouldn't be blown off course by bad weather and we are not going to be.

Pym Won't this add weight to Ed Ball's argument that by embarking on these cuts you are putting growth at risk?

Osborne Well, if you look at the December period, with the very bad weather, the worst weather for a hundred years, of course actually the tax rises and the spending review process had not kicked in then, and so that is not an excuse that people can make. We are very clear that to abandon the budget plans, as the Labour Party would have us do, would put us back into the financial crisis zone, which is where the Labour Party left us. We are not going to do that. We are not going to be blown off course by bad weather. The economy needs to rebalance and you see manufacturing growing at the moment.

Pym Isn't there every chance that this quarter, the first quarter of 2011, there'll be persistent weaknesses, partly because of the VAT rise?

Osborne Well, as I say, we got these figures today. They are very uncertain, and the impact of the weather has clearly been enormous, as the office of national statistics, who put together the forecast, has made very clear. And it was the coldest December for a hundred years, people couldn't get to work, businesses were closed, and that has had a bigger impact than anyone forecast. But if you look at areas not so affected by the weather, like manufacturing, they are growing. That is an important part of rebalancing the British economy, and if we were to abandon our budget plans, and not face up to the debts, as the way that Labour suggests, then we would be back in a financial crisis. That would be a disaster for Britain, and this Government is not going to be blown off course by bad weather.

Pym Can I ask you one question about the talks with the banks, as the final one? I mean, are you close to an agreement with the banks on lending and bonuses and so on?

Osborne Well, we are engaged in a conversation with the banks. I've made that very clear. What we want to see is more lending, we want to see small bonuses, and we want to see the banks paying more taxes; and that's what I hope we can achieve. That'd be good for the British economy, good for the British taxpayer and actually also good for British financial services, which employs hundreds of thousands of people.

Pym Are you nearly there with those talks?

Osborne Well, we are having those conversations and I hope we can reach a settlement, but we've set out the terms of that settlement very clearly.

Based on your intuitions on how participants *ought to behave* in a political interview, how do you rate their performance in this fragment?

	Incorrect	Mostly incorrect	Somewhere in the middle	Mostly correct	Correct
Pym	○	○	○	○	○
Osborne	○	○	○	○	○

Please answer the following questions about your familiarity with the interview.

	Yes	No	Not Sure
Have you watched or read this interview before?	○	○	○
Are you familiar with the political/historical context?	○	○	○
Have you heard of the interviewer?	○	○	○
Have you heard of the interviewee?	○	○	○

9. Conclusion Exit this survey

All done!

Thank you for completing this survey. Please answer the following question.

Have you read the interview fragments you responded to in detail?

- ◯ Yes
- ◯ No

If you have any comments, please leave them in the box below.

And here you can watch some of the interviews I have analysed as part of my research. Enjoy!

(Source: BBC NEWS)

(Source: BBC NEWS)

(Source: BBC NEWS)

C.2 Dissemination

The survey was disseminated via email to the mailing lists of four research special interest groups[2], the Postgraduate Student Forum at the Open University Computing Department and via a series of posts on the social networking site Facebook[3]. Volunteers were invited to share the post with their contacts on the site.

C.2.1 Email Message Sent to Research Email Lists

```
(Apologies for cross-postings)

Dear SIGDIAL/SIGSEM/CLUK/ELSNET members,

If you can spare a few minutes, please help our research by completing this
survey on how people perceive verbal behaviour in political interviews:

https://www.surveymonkey.com/s/2ZTCT6B

The entire survey should take about 20 minutes to complete, *but* if you do
not have that amount of time to spare, please complete as much of it as you
can and then skip through the rest until you get to the final page

There is a treat at the end of the survey, where you can watch a few rather
amusing interactions we have come across during our research.

Thanks for your help!

------------
Brian Pluss
Natural Language Generation Group
Computing Department
The Open University
Walton Hall
Milton Keynes
MK7 6AA
United Kingdom
```

[2]SIGIAL (http://www.sigdial.org/), SIGSEM (http://www.sigsem.org/), CLUK (http://nlp.shef.ac.uk/research/cluk/) and ELSNET (http://www.elsnet.org/).

[3]https://www.facebook.com/

C.2.2 Facebook Post

```
Dear friends, this is serious business...

If you are fluent in English, please (please!) help my research by
completing this survey on how people perceive verbal behaviour in
political interviews:

https://www.surveymonkey.com/s/2ZTCT6B

Feel free to share the link above and spread the word... This is
one of those the-more-the-merrier kind of things.

Thank you!

PS: there's small a treat at the end of the survey.
```

Appendix D

A Discussion on Modelling Turn Length

We discuss here an idea on how to model the number of contributions agents make in each turn. As described in Section 5.1.8, when deciding on their next dialogue move, agents use a biased random function to choose whether they actually make a move or whether they release the floor. Each agent has a a parameter q to influence the likelihood of releasing the turn each time they select a next dialogue move. However, using a Bernoulli trial with probability of success q, as we do in the deliberation mechanism, could lead to agents releasing the floor immediately after taking it. To avoid this, we take a different approach, making the probability of success change along the turn.

We propose using an *exponential decay function* to model a decrease in likelihood that an agent keeps the floor after each move in a same turn. This is a function $q(k) = e^{-qk}$, where k is the number of moves in the current turn (the first move being 0) and q is the decay rate parameter chosen for the agent. The higher the value for q the steeper the decay, so values closer to 0 will make it more likely that the agent performs more moves after the initial one. Figures D.1 to D.4 show this function for different values of q.

For the function $release\text{-}floor$ we would use a Bernoulli trial with probability of success $r(k) = 1 - q(k) = 1 - e^{-qk}$. This means that in the first move of a turn, the trial will succeed with probability $r(0) = 1 - q(0) = 1 - e^{-q0} = 0$, regardless of the parameter q chosen for the agent, meaning that the function will fail and the agent will select a move. For the second move of the turn, the probability of success would be $r(1) = 1 - e^{-q}$. For $q = 0.2$, it is $r(1) \approx 0.18$; for $q = 0.8$ it is $r(1) \approx 0.55$; for $q = 2$ it is $r(1) \approx 0.86$; and for $q = 5$ it is $r(1) \approx 0.99$. This means that, the higher the value of $q \in [0, 5]$ the more likely it is that the agent releases the floor after the first move. In the second move within a turn, for $q = 0.2$, it is $r(2) \approx 0.33$, while for $q = 2$ it is $r(2) \approx 0.98$. Figure D.5 shows $r(k)$ for these four values of q.

In order to determine the value of q for each agent, we can look at the distribution of turn length over n trials of the $release\text{-}floor$ function. This is like a binomial distribution of failures of the function (meaning that the floor is not released) over n attempts. Using the probability mass function for a binomial distribution[1] but using $q(k)$ in the place of p, we get the function:

$$f(k; n, q) = \binom{n}{k} q(k)^k \left(1 - q(k)\right)^{n-k} = \binom{n}{k} \left(e^{-qk}\right)^k \left(1 - e^{-qk}\right)^{n-k}$$

A graphic of this function for the four values of q above and $n = 10$ is shown in Figure D.6.

[1] The probability mass function for the binomial distribution with parameters n and k is $f(k; n, p) = Pr(K = k) = \binom{n}{k} p^k (1-p)^{n-k}$, with $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

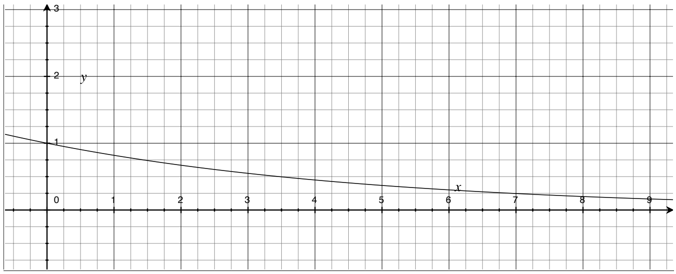


Figure D.1: Exponential decay function for decay rate $q = 0.2$.

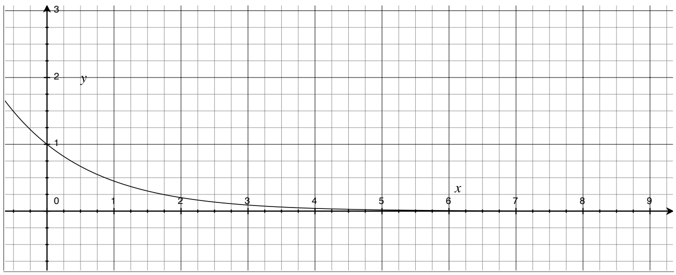


Figure D.2: Exponential decay function for decay rate $q = 0.8$.

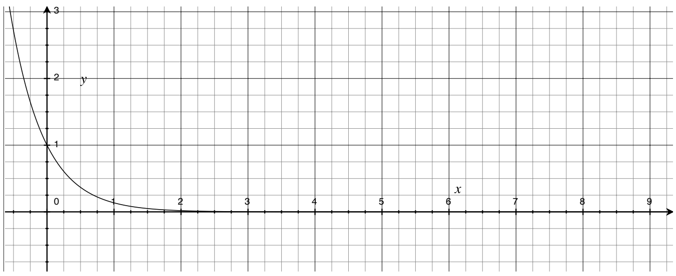


Figure D.3: Exponential decay function for decay rate $q = 2$.

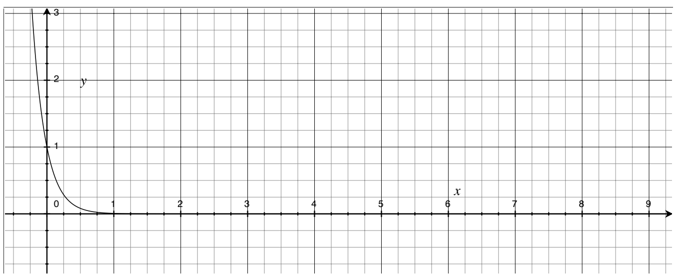


Figure D.4: Exponential decay function for decay rate $q = 5$.

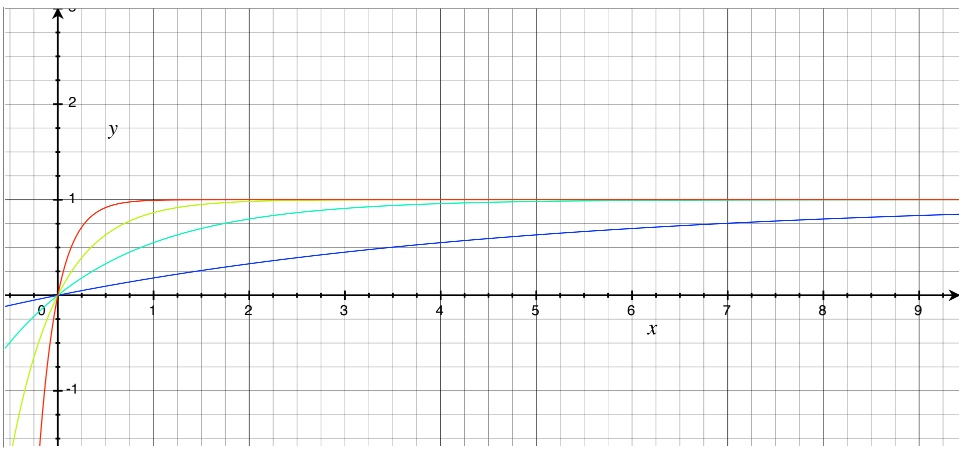


Figure D.5: Probability of success $r(k)$ for $q \in \{0.2, 0.8, 2, 5\}$.

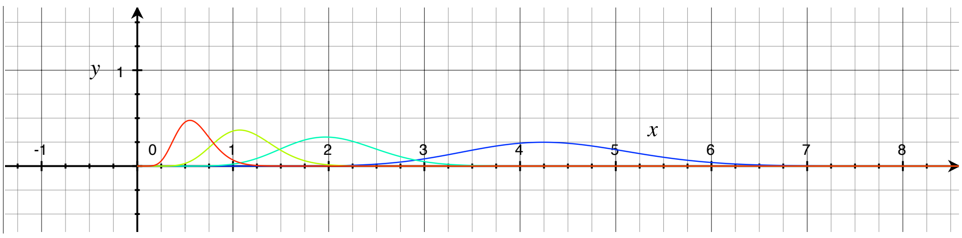


Figure D.6: Distribution of turn length over 10 trials of the *release-floor* function for $q \in \{0.2, 0.8, 2, 5\}$.

Appendix E

Computing Dynamic Obligations for D_1

This appendix shows the update of dynamic obligations for the (hand-crafted) political interview in Figure E.1. The example was first introduced in Figure 3.7 and used in Chapter 5 to illustrate the proposed model of

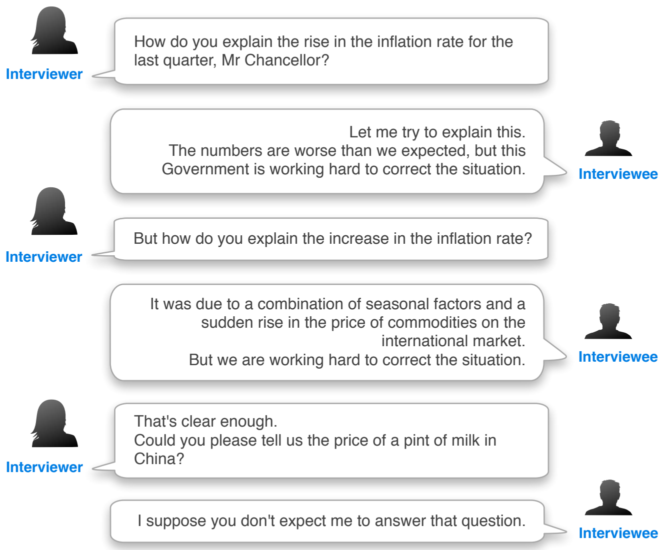


Figure E.1: A hand-crafed political interview transcript conversational agents. Recall the formalisation in terms of action labels given in Chapter 3:

$$\begin{aligned}D_1 = \langle &(\mathtt{ir}, \langle(1) : \mathtt{valid\text{-}question}\ \mathbf{N}\rangle);\\ &(\mathtt{ie}, \langle(2) : \mathtt{acceptance@}(1); (3) : \mathtt{invalid\text{-}reply@}(1)\ \mathbf{I}\rangle);\\ &(\mathtt{ir}, \langle(4) : \mathtt{valid\text{-}question}\ \mathbf{R}\rangle);\\ &(\mathtt{ie}, \langle(5) : \mathtt{valid\text{-}reply@}(1)\ \mathbf{C}; (6) : \mathtt{invalid\text{-}reply@}(1)\ \mathbf{C}\rangle);\\ &(\mathtt{ir}, \langle(7) : \mathtt{acceptance@}(5); (8) : \mathtt{invalid\text{-}question}\ \mathbf{N}\rangle);\\ &(\mathtt{ie}, \langle(9) : \mathtt{rejection@}(8)\rangle)\rangle\end{aligned}$$

and the dialogue game: $G_{PI} = (Allow_{PI}, Introduce_{PI}, Discharge_{PI})$, with:

$$\begin{aligned}
Allow_{PI} = \{&[\mathtt{ir} : \{\mathtt{valid\text{-}statement}, \mathtt{valid\text{-}question}, \mathtt{acceptance}, \mathtt{rejection}\}], && (1)\\
&[\mathtt{ie} : \{\mathtt{valid\text{-}statement}, \mathtt{valid\text{-}reply}, \mathtt{acceptance}, \mathtt{rejection}\}]\} && (2)\\
Introduce_{PI} = \{&[(\mathtt{ir}, (\mathtt{s}) : \mathtt{valid\text{-}statement}) \rightsquigarrow (\mathtt{ie}, \mathtt{acceptance@}(\mathtt{s}))], && (3)\\
&[(\mathtt{ir}, (\mathtt{q}) : \mathtt{valid\text{-}question}\ \mathbf{N}) \rightsquigarrow (\mathtt{ie}, \mathtt{acceptance@}(\mathtt{q}))], && (4)\\
&[(\mathtt{ie}, \mathtt{acceptance@}(\mathtt{q})) \rightsquigarrow (\mathtt{ie}, \mathtt{valid\text{-}reply@}(\mathtt{q})\ \mathbf{C})], && (5)\\
&[(\mathtt{ie}, (\mathtt{s}) : \mathtt{valid\text{-}statement}) \rightsquigarrow (\mathtt{ir}, \mathtt{acceptance@}(\mathtt{s}))], && (6)\\
&[(\mathtt{ie}, (\mathtt{r}) : \mathtt{valid\text{-}reply@}(\mathtt{q})) \rightsquigarrow (\mathtt{ir}, \mathtt{acceptance@}(\mathtt{r}))], && (7)\\
&[(\mathtt{ir}, \mathtt{acceptance}) \rightsquigarrow (\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})], && (8)\\
&[(\mathtt{ir}, (\mathtt{s}) : \mathtt{invalid\text{-}statement}) \rightsquigarrow (\mathtt{ie}, \mathtt{rejection@}(\mathtt{s}))], && (9)\\
&[(\mathtt{ir}, (\mathtt{q}) : \mathtt{invalid\text{-}question}) \rightsquigarrow (\mathtt{ie}, \mathtt{rejection@}(\mathtt{q}))], && (10)\\
&[(\mathtt{ir}, (\mathtt{r}) : \mathtt{invalid\text{-}reply}) \rightsquigarrow (\mathtt{ie}, \mathtt{rejection@}(\mathtt{r}))], && (11)\\
&[(\mathtt{ie}, (\mathtt{s}) : \mathtt{invalid\text{-}statement}) \rightsquigarrow (\mathtt{ir}, \mathtt{rejection@}(\mathtt{s}))], && (12)\\
&[(\mathtt{ie}, (\mathtt{q}) : \mathtt{invalid\text{-}question}) \rightsquigarrow (\mathtt{ir}, \mathtt{rejection@}(\mathtt{q}))], && (13)\\
&[(\mathtt{ie}, (\mathtt{r}) : \mathtt{invalid\text{-}reply}) \rightsquigarrow (\mathtt{ir}, \mathtt{rejection@}(\mathtt{r}))]\} && (14)\\
Discharge_{PI} = \{&[*\text{-}\mathtt{question}\ \mathbf{R} \succ \mathtt{rejection}], && (15)\\
&[*\text{-}\mathtt{statement} \succ \mathtt{acceptance}], && (16)\\
&[*\text{-}\mathtt{question}\ \mathbf{N} \succ \mathtt{acceptance}], && (17)\\
&[*\text{-}\mathtt{reply} \succ \mathtt{acceptance}]\} && (18)
\end{aligned}$$

The sequence of dynamic obligations for each turn is computed as shown below[1]:

$$\begin{aligned}
O_{D_1} &= \langle\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle\rangle && (0)\\
\xrightarrow{\text{Turn 1}} &\quad [\![(\mathtt{ir}, (1) : \mathtt{valid\text{-}question}\ \mathbf{N}) + \text{discharge}]\!]\\
O_{D_1} &= \langle\langle(\mathtt{ir}, \mathtt{valid\text{-}question})\rangle; && (0)\\
&\quad \langle\rangle\rangle && (1)
\end{aligned}$$

[1]The reasons for each change are given in double square brackets next to the arrows labelled with the turn in which they occur. The numbers on the right indicate the turn to which the sequence of obligations corresponds.

$$\xrightarrow{\text{Turn 1}} \quad [\![(\mathtt{ir}, (1) : \mathtt{valid\text{-}question}\ \mathbf{N}) + \text{introduction Rule (4)}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle\rangle && (1) \end{aligned}$$

$$\xrightarrow{\text{Turn 2}} \quad [\![(\mathtt{ie}, (2) : \mathtt{acceptance@}(1)) + \text{discharge}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle\rangle\rangle && (2) \end{aligned}$$

$$\xrightarrow{\text{Turn 2}} \quad [\![(\mathtt{ie}, (2) : \mathtt{acceptance@}(1)) + \text{introduction Rule (5)}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle\rangle && (2) \end{aligned}$$

$$\xrightarrow{\text{Turn 2}} \quad [\![(\mathtt{ie}, (3) : \mathtt{invalid\text{-}reply@}(1)\ \mathbf{I}) + \text{introduction Rule (14)}$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle(\mathtt{ir}, \mathtt{rejection@}(3)); (\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle\rangle && (2) \end{aligned}$$

$$\xrightarrow{\text{Turn 3}} \quad [\![(\mathtt{ir}, (4) : \mathtt{valid\text{-}question@}(3)\ \mathbf{R}) + \text{discharge Rule (15)}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle(\mathtt{ir}, \mathtt{rejection@}(3)); (\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (2)\\ &\langle(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle\rangle && (3) \end{aligned}$$

$$\xrightarrow{\text{Turn 4}} \quad [\![(\mathtt{ie}, (5) : \mathtt{valid\text{-}reply@}(1)\ \mathbf{C}) + \text{discharge}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle(\mathtt{ir}, \mathtt{rejection@}(3)); (\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (2)\\ &\langle(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (3)\\ &\langle\rangle\rangle && (4) \end{aligned}$$

$$\xrightarrow{\text{Turn 4}} \quad [\![(\mathtt{ie}, (5) : \mathtt{valid\text{-}reply@}(1)\ \mathbf{C}) + \text{introduction Rule (7)}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle(\mathtt{ir}, \mathtt{rejection@}(3)); (\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (2)\\ &\langle(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (3)\\ &\langle(\mathtt{ir}, \mathtt{acceptance@}(5))\rangle\rangle && (4) \end{aligned}$$

$$\xrightarrow{\text{Turn 4}} \quad [\![(\mathtt{ie}, (6) : \mathtt{invalid\text{-}reply}\ \mathbf{C}) + \text{introduction Rule (14)}]\!]$$

$$\begin{aligned} O_{D_1} = \langle &\langle(\mathtt{ir}, \mathtt{valid\text{-}question}\ \mathbf{N})\rangle; && (0)\\ &\langle(\mathtt{ie}, \mathtt{acceptance@}(1))\rangle; && (1)\\ &\langle(\mathtt{ir}, \mathtt{rejection@}(3)); (\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (2)\\ &\langle(\mathtt{ie}, \mathtt{valid\text{-}reply@}(1)\ \mathbf{C})\rangle; && (3)\\ &\langle(\mathtt{ir}, \mathtt{rejection@}(6)); (\mathtt{ir}, \mathtt{acceptance@}(5))\rangle\rangle && (4) \end{aligned}$$

$$\xrightarrow{\text{Turn 5}} \quad [\![(\mathtt{ir}, (7) : \mathtt{acceptance@}(5)) + \text{discharge}]\!]$$

$$
\begin{aligned}
O_{D_1} = \langle & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle; && (0)\\
& \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle; && (1)\\
& \langle(\texttt{ir}, \texttt{rejection@}(3)); (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (2)\\
& \langle(\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (3)\\
& \langle(\texttt{ir}, \texttt{rejection@}(6)); (\texttt{ir}, \texttt{acceptance@}(5))\rangle; && (4)\\
& \langle(\texttt{ir}, \texttt{rejection@}(6))\rangle\rangle && (5)
\end{aligned}
$$

$$\xrightarrow{\text{Turn 5}} \quad [\![(\texttt{ir}, (7) : \texttt{acceptance}) + \text{introduction Rule (8)}]\!]$$

$$
\begin{aligned}
O_{D_1} = \langle & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle; && (0)\\
& \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle; && (1)\\
& \langle(\texttt{ir}, \texttt{rejection@}(3)); (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (2)\\
& \langle(\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (3)\\
& \langle(\texttt{ir}, \texttt{rejection@}(6)); (\texttt{ir}, \texttt{acceptance@}(5))\rangle; && (4)\\
& \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N}); (\texttt{ir}, \texttt{rejection@}(6))\rangle\rangle && (5)
\end{aligned}
$$

$$\xrightarrow{\text{Turn 5}} \quad [\![(\texttt{ir}, (8) : \texttt{invalid-question}\ \mathbf{N}) + \text{introduction Rule (10)}]\!]$$

$$
\begin{aligned}
O_{D_1} = \langle & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle; && (0)\\
& \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle; && (1)\\
& \langle(\texttt{ir}, \texttt{rejection@}(3)); (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (2)\\
& \langle(\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (3)\\
& \langle(\texttt{ir}, \texttt{rejection@}(6)); (\texttt{ir}, \texttt{acceptance@}(5))\rangle; && (4)\\
& \langle(\texttt{ie}, \texttt{rejection@}(8)); (\texttt{ir}, \texttt{valid-question}\ \mathbf{N}); (\texttt{ir}, \texttt{rejection@}(6))\rangle\rangle && (5)
\end{aligned}
$$

$$\xrightarrow{\text{Turn 6}} \quad [\![(\texttt{ie}, (9) : \texttt{rejection@}(8)) + \text{discharge}]\!]$$

$$
\begin{aligned}
O_{D_1} = \langle & \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N})\rangle; && (0)\\
& \langle(\texttt{ie}, \texttt{acceptance@}(1))\rangle; && (1)\\
& \langle(\texttt{ir}, \texttt{rejection@}(3)); (\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle && (2)\\
& \langle(\texttt{ie}, \texttt{valid-reply@}(1)\ \mathbf{C})\rangle; && (3)\\
& \langle(\texttt{ir}, \texttt{rejection@}(6)); (\texttt{ir}, \texttt{acceptance@}(5))\rangle; && (4)\\
& \langle(\texttt{ie}, \texttt{rejection@}(8)); (\texttt{ir}, \texttt{valid-question}\ \mathbf{N}); (\texttt{ir}, \texttt{rejection@}(6))\rangle; && (5)\\
& \langle(\texttt{ir}, \texttt{valid-question}\ \mathbf{N}); (\texttt{ir}, \texttt{rejection@}(6))\rangle\rangle && (6)
\end{aligned}
$$

Index