Citation:
Adam, C, Cavedon, L and Padgham, L 2010, '"Hello Emily, how are you today?" Personalised dialogue in a toy to engage children', in Yorick Wilks (ed.) Proceedings of the 2010 Workshop on Companionable Dialogue Systems, ACL 2010, United States, 15 July 2010, pp. 19-24.

"Hello Emily, how are you today?" Personalised dialogue in a toy to engage children

Carole Adam
RMIT University
Melbourne, Australia.
carole.adam.rmit@gmail.com

Lawrence Cavedon
RMIT University
Melbourne, Australia.
lawrence.cavedon@rmit.edu.au

Lin Padgham
RMIT University
Melbourne, Australia.
lin.padgham@rmit.edu.au

Abstract

In line with the growing interest in conversational agents as *companions*, we are developing a toy companion for children that is capable of engaging interactions and of developing a long-term relationship with them, and is extensible so as to evolve with them. In this paper, we investigate the importance of personalising interaction both for engagement and for long-term relationship development. In particular, we propose a framework for representing, gathering and using personal knowledge about the child during dialogue interaction. [1]

1 Introduction

In recent years there has been an increasing interest in so-called *Companion* agents: agents that are intelligent, and built to interact naturally (via speech and other modalities) with their user over a prolonged period of time, personalising the interaction to them and developing a relationship with them. The EU Companions project[2] is the most well known such project, with applications such as a companion for the elderly (Field et al., 2009), and a health and fitness companion (Stahl et al., 2009). In our work, together with industry partners, we are developing a speech-enabled companion toy for children. While there are many "smart toys" on the market, as far as we are aware our work is unique in attempting to develop a "companion toy" for a child, evolving with them over a long period of time. As with other projects on intelligent companions, a crucial task is to build a long-term relationship with the user, by a series of interactions over time, that the user experiences as engaging and valuable.

According to models of the "enjoyability" of human-computer interaction (Brandtzaeg et al., 2006), there are three main features making an interactive system engaging for the user: the user should feel in **control** of the interaction (which includes being able to customise it and getting timely feedback); the **demands** on the user should be adapted to their capabilities, *i.e.* the interaction should be challenging and surprising but not overwhelming; and the system should **support** social interaction rather than isolating the user. Another important aspect of any engaging interaction is for it to be **personalised**, *i.e.* customised to the particular interlocutor and their environment. Other important features for engagement include coherence of the dialogue, emotional management, and personality. In this paper we focus specifically on the issue of appropriate personalisation of interactions with a child, and how to realise this.

Existing personalised systems mainly have a *task-oriented* focus, *i.e.* they aim at building a user profile and using it to facilitate the user's task (*e.g.* Web navigation assistants or product recommendation systems (Abbattista et al., 2003)), and at being user-configurable. On the contrary we aim at personalising the interaction to build a relationship and engage a child. The main novelties of our system are that: it is *not* task-oriented; it is specifically designed for children; and its behaviour is derived from actual interaction data. Indeed, in order to understand the kinds of personalisation occurring in natural dialogues with children, we have analysed corpora of children's dialogues (MacWhinney, 1995; MacWhinney, 2000). We have then developed a framework that enables the implementation of a number of these personalised behaviours within our intelligent toy.

The contribution of this paper is the **identification** of different kinds of personalisation behaviours in dialogue with children, based on actual data, plus the **framework** to realise these within an implemented system.

[1] A slightly longer version of this paper is currently under review elsewhere. If both papers are accepted for publication we will modify to ensure that they expand different aspects.

[2] See www.companions-project.org.

Proceedings of the 2010 Workshop on Companionable Dialogue Systems, ACL 2010, pages 19–24,
Uppsala, Sweden, 15 July 2010.

2 Personalisation behaviours

2.1 Corpus analysis

We have analysed examples of children-adult dialogues (mainly from the CHILDES database (MacWhinney, 1995; MacWhinney, 2000); one dialogue from a forthcoming study performed with a puppet as part of this project) in order to determine the types of behaviours that adults use to personalise their interaction with a child.

Relation to self

A first observation is that children often try to relate conversation to themselves. This is illustrated by this conversation between a girl (G) and her mother (M) about a visit to the doctor.

G What's polio?
M An illness that makes you crippled. That's why you get all those injections and... A long time ago, kiddies, kiddies used to die with all that things.
G will I ?
M hmm. You aren't going to die.

Personal questions

Adults also often ask the child questions about themselves. This dialogue illustrates a conversation between an adult (A) and a child (C) about C's holidays. Notice that the questions are adapted to the context (ask about holidays in summer).

A Did you go on vacation over the summer? Did you?
A Where'd you go? To the beach?
C Yes.
A Yeah? Did you go by yourself? No. Why laugh? You could go by yourself.
A Do you have brothers and sisters?
C Just a little sister.
A A sister? Did she go too? On vacation?

Child control

Even if the adult is asking the questions, the child retains some control over the interaction. The following dialogue between a boy (B) and his grandmother (G) shows how the adult follows the child when he switches away from a disliked topic. This dialogue also shows the adult commenting on the child's tastes based on her knowledge of them.

G how are you getting on in school?
B we're not going to go shopping today.
G eh?
B shopping today.
G ...
B and chips.
G going to have chips?
B mm.
G you likes that.

Reciprocity

Another way for the adult to learn personal information about the child without asking questions is to confide personal information first, which encourages the child to reciprocate. In this dialogue between a child (C) and a puppet (P) controlled by an adult, P confides personal information (its tastes), which leads the child to do the same.

P My favourite drink is lemon. Lemon soft drink. I like that.
C Mine is orange juice.
P mmhm. Orange one? You like the orange one?
*C Orange juice (*nodding*)*

Recalling shared activities

Another form of personalisation is recalling past shared activities. In the following dialogue, a mother (M) reads a book to her child (C); when a picture of a snowman appears in the book she recalls the child recently making one with her.

M what did we make outside here today?
C um I don't know.
M did we make a man?
C yeah.
M a snowman?
C yeah.

Child's preferences

Another way to personalise interaction is to recall a child's preferences. For example this dialogue involves a child (C) and an interrogator (I) wanting to record a story. Here the child corrects incorrect knowledge; this update should be remembered.

I Do you wanna tell a story?
C No. I won't.
I No, you don't.
I You told me down there that you like stories.
C No, I hate stories.

Child's agenda

Parents may also use knowledge about a child's agenda (*i.e.* planned future activities, school, etc.) and make relevant and timely comments about it. In this dialogue a mother (M) and her friend (F) talk with a boy (B) about his next school day, when he is supposed to see chicken eggs hatching.

F Oh you're going to see the little chicks tomorrow are you. You'll have to tell me what it's like. I haven't never seen any.
B I I haven't either.
F I haven't.
M We've seen them on the tellie, haven't we?
F I haven't seen those little ones.
M haven't you?
F So you'll have to tell me.
M Have you seen them on the tellie?
B mm [= yes].

We notice again that when the mother's friend confides some information (she never saw that), the child reciprocates (he neither). Moreover the mother again shows memory of past activities (seeing something on television).

2.2 Personalisation strategies

Based on our analysis of adult-children dialogue corpora, we have designed a number of strategies to allow our toy to generate these kinds of personalised interactions with the child. These strategies fit into two categories: strategies for **gathering** personal information, and strategies for **exploiting** personal information.

Information gathering

The Toy can gather and then use different types of information: (1) personal information (e.g. family, friends, pets); (2) preferences (e.g. favourite movie, favourite food); (3) agenda (plays football on Saturday, has maths every Thursday); (4) activity-specific information (preferred stories, current level of quiz difficulty); (5) interaction environment (e.g. time, day, season, weather).

The easiest strategy to gather this information is to explicitly query the child. These queries have to be made opportunistically, *e.g.* when matching the current conversational topic, so as to seamlessly integrate information gathering into a conversation. Other strategies include confiding personal information to make the child reciprocate and confide similar information; or extracting personal information from spontaneous child's input. These strategies are useful so as to avoid asking too many questions, which would dirupt the conversation flow and could annoy the child.

Information exploitation

One of the challenges for using the gathered personal information in a conversation is to determine the appropriate opportunities to do so. The personal information can be used to engage the child in various ways, reproducing the types of behaviours illustrated above. In particular, our toy has the following information exploiting strategies: (1) use child's name; (2) insert comments using personal information; (3) ask about daily activities; (4) adapt interaction (*e.g.* greetings) to the context (*e.g.* time of day); (5) take child's preferences into account in topic or activity selection.

3 The Toy architecture: overview

This section outlines the general architecture of the toy. The integration of our personalisation framework is detailed in Section 4.

The central component of the Toy is the *Dialogue Manager* (DM) which is made up of two components: the *input/output manager* (IOM) receives input from Automatic Speech Recognition (ASR)[3] and sends output to Text-to-Speech (TTS); the *Semantic Interaction Manager* (SIM) receives input from IOM, generates the toy's response and sends it back to IOM (see Figure 1).

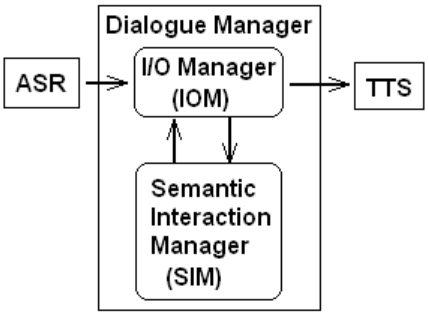

Figure 1: Architecture of the Toy

Our current approach to ASR and utterance processing is grammar-based: on sending an output utterance for synthesis, the DM loads into the speech recogniser a parameterised grammar specifying the set of expected user responses to this output. The DM is multi-domain and extensible via *domain modules*, designed to handle utterances about a particular domain, and encapsulating data required for this: a knowledge-base segment; a set of *conversational fragments* (see Section 3.2.2); a collection of the *topics* it is designed to handle; and an entry grammar to assign a topic to inputs.

3.1 Input Output Manager

The IOM is implemented using a BDI agent-oriented methodology, with dialogue processing "strategies" built as plans. For example, there are plans designed to handle errors or low-confidence results from speech recognition; plans to handle utterance content and update the information state; and plans to manage concurrent conversational threads and select which of a number of candidate responses to output.

3.2 Semantic Interaction Manager

The Semantic Interaction Manager (SIM) is a component designed to manage flexible conversational flow. The SIM maintains an *agenda* of things to say. When an input is received from the IOM, it is pre-processed to generate an input analysis that informs the further stages of the

[3]We have mainly used SRI's *Dynaspeak* system which is designed for small computational platforms.

SIM plan. In particular the input is then either dispatched to an existing ongoing *activity* if it matches its expected answers, or an appropriate new activity is created. The chosen activity selects a *conversational fragment* in the topic network corresponding to its topic, and writes it in the conversational agenda. Finally the output is generated from the agenda and sent to the IOM.

3.2.1 The conversational agenda

The conversational agenda maintained by the SIM has two main parts. The *history* represents the past interaction and stores past *questions under discussion* (QUD) (Ginzburg, 1997) with their received answer. The *stack* represents the future interaction and lists QUD to be asked next, in order. The agenda also stores the current ongoing activities (Section 3.2.3), making it possible to switch back and forth between them.

3.2.2 Conversational fragments

In our system, we use pre-scripted pieces of dialogue that we call *conversational fragments*. The designers of domain modules will provide a *topic network* describing its domain, with nodes being the possible topics, having links with other topics, and providing a pool of fragments to possibly use when talking about this topic. Each fragment has an applicability condition, and provides the text of an output as well as a list of expected answer patterns with associated processing (*e.g.* giving feedback) applied when the child's response matches.

This representation obviates the need for full natural language generation (NLG) by providing semi-scripted outputs, and also informs the grammar-based ASR by providing a list of expected child answers. Moreover it allows the Toy to generate quite flexible interactions by switching between topics and using fragments in any order.

3.2.3 Activities

When interacting with the child, the Toy suggests possible *activities* (*e.g.* quiz, story) about the available topics. Each type of activity uses specific types of fragments (*e.g.* quiz questions with expected (in)correct answers; story steps with expected questions) and has particular success and failure conditions (*e.g.* a number of (in)correct answers for a quiz; or reaching the end for a story).

This concept of activity helps to keep the dialogue cohesive, while allowing flexibility. It also meets the requirement that an engaging interaction should be *demanding* for the child while staying *controlled* by them. Indeed a number of activities can be listed in the agenda at the same time, being resumed or paused to allow switching between them (*e.g.* to follow the child's topic requests or to insert personalised contributions).

4 The toy personalisation framework

We now describe our framework for implementing the personalisation strategies specified earlier.

4.1 The personalisation frame

All the information that our toy needs to personalise an interaction is gathered using a structure called the *personalisation frame*. This structure is tailored to the requirements imposed by our architecture, namely the grammar-based speech recognition and the absence of natural language processing. It consists of: (1) a static list of personal information **fields** (*e.g.* child name, age); (2) a static indexed list of **rules** specifying when it is appropriate to insert personal comments or questions in the interaction; (3) a dynamic child **profile**, storing the current values of (some) personal information fields, updated during interaction.

Personal information fields (PIFs)

Each personal information field contains: a list of possible values for this field (informing the ASR grammar); and a grammar of specific ways in which the child may spontaneously provide information relevant to this field (allowing the toy to interpret such input and extract the value).

For example the field "favourite animal" has a list of animals as its values, and its grammar contains patterns such as "My favourite animal is X" or "I love X" (where the variable X ranges over the possible values of this field).

Personalisation rules

Each personalisation rule specifies the opportunity that triggers it, and the text of the output. The text of personalisation comments and questions is scripted, and used to automatically generate conversation fragments from the frame. **Comment rules** also specify the list of personal information fields that are used in the text of the comment, while **Question rules** specify the name of the field set by their answer and a grammar of expected answers, with their interpretation in terms of which value the corresponding field should receive.

For example, there may be a *comment rule* referring to the field pet_type, enabling the output "I know you have a pet_type" when the keyword pet_type is detected. There may also be a *question rule* for asking "What is your favourite animal?" when talking about the zoo; expected answers would include "I like A"; so if the child answers "I like tigers" then the $favourite\_animal$ field would receive the value "tigers" as a result.

Opportunities

Personalisation must be integrated into the conversational management so as not to disrupt dialogue (*i.e.* the toy should still maintain a coherent interaction). It is thus important to accurately detect appropriate opportunities to insert personalisation side-talk. There are three types of opportunities that can trigger the personalisation rules: (1) **keyword opportunities** (a particular keyword appears in the child's input, *e.g.* the child uses the word "mother"); (2) **topic opportunities** (the interaction is focused on a particular topic, *e.g.* the child is talking about koalas); (3) **activity opportunities** (a particular activity is in a particular state, *e.g.* start of a story).

The following sections describe how this *personalisation frame* is used in the Conversation Manager process to personalise the conversation that is generated: we first outline the full process, before giving details about the steps where the *personalisation frame* is used.

4.2 Personalised input handling

The following algorithm is the result of the integration of personalisation into the response generation plan of the SIM. Steps manipulating the personalisation frame will be detailed below.

1. Initialisation (load child profile, update environment description);
2. Input reception (from IOM):
3. Input analysis (preprocess input, detect opportunities);
4. Profile update;
5. Input dispatching (to selected activity);
6. Activity progressing (fragment selection);
7. Personalisation generation (generate fragment from best applicable triggered rule);
8. Agenda processing (prioritisation of activity vs personalisation fragments);
9. Personalisation of output (detection of opportunities, modification of output);
10. Output generation (sent to IOM);
11. End turn (save profile).

Fragment selection (step 6)

Fragment selection is personalised in two ways. **First**, some fragments have applicability conditions concerning the interaction context and the child's profile. For example a fragment such as "Hi, what's your name?" is only applicable if the toy does not know the child's name. A greeting fragment such as "Hi! How was school today?" is only applicable at the end of a school day. Other greeting fragments are available for different contexts. **Second**, some fragments have an adaptable content, using variables referring to the child's profile and to the context. These fragments are only applicable if the value of these variables is known and can be used to instantiate the variable when generating output. For example a fragment with the text "Hello $child\_name$! How are you?" is applicable once the child's name is known. Or a fragment saying "I know you have a pet_type called pet_name." will be instantiated as "I know you have a cat called Simba".

Personalisation fragments generation (step 7)

When an opportunistic rule in the personalisation frame is triggered, its applicability is checked: *comment rules* are only applicable if the fields used have a value; *question rules* are only applicable if the field set has no value. Then the applicable rule of highest priority is used to generate a personalisation fragment. Its topic is the current topic provided in the input analysis; its type is either "personal question" or "personal comment"; and its text is as specified in the rule. Comment fragments have no additional applicability condition and no expected answers. Question fragments receive the generic expected answers specified in the rule, instantiated with the possible values of the field set by this question; the process associated with them consists in setting the value of this field to the value extracted from the input; the default process stores the full input as the value.

For example with the *question rule* "What is your favourite animal?" shown above, the generated fragment will expect answers such as "I like tigers", "I like giraffes", *etc*. If a matching answer is received, then the corresponding animal is stored as the child's favourite one in their profile.

Prioritisation (step 8)

When updating the agenda, several situations can arise depending on the relative priorities of personalising the interaction vs continuing the current ac-

tivity. If the activity can be disrupted at this point, personalisation question fragments can be added on top of the stack in the agenda, while comments can be appended to the activity output; if the personalisation fragment has too low priority to disrupt the activity, then it is dropped. This priority of personalisation depends in particular on the priority of the triggered rule and on the time since personalisation was last inserted.

5 Conclusion

This framework for personalisation has been implemented in our toy using the Java-based agent platform Gorite (Ronnquist, 2008). The following is an example of the type of dialogues we intend to generate with it (T is the Toy and C a child); it illustrates a number of our personalisation strategies. Our infrastructure technically allows it to happen, however we still need to add more content (domain modules and fragments in them) to make it actually happen; we are working on writing such content, as well as on providing tools to make it easier for designers to write this content in the future.

C Hi there
T Hello Emily! `[Use child's name]`
T What did you do on such a sunny week-end? `[Context-adaptation]`
C I went to the zoo!
T Great! That must have been fun. Which animals did you see?
C I saw tigers
T Tigers are cats, but much bigger than your cat Simba. What else did you see?
C Koalas!
T Oh, that's your favourite animal, isn't it? `[Personal comment]`
C Yes
T You must have been really happy. My favourite animal is orangutan, did you see any? `[Confide personal information]`
C No
T Pity... Allright, enough chatting. I know you love stories. Do you want to hear a story about tigers? `[Choose favourite activity]`

The ASR is not functional yet due to the specific challenges of recognizing children voices, so for now we are only experimenting with a textual interface. This may look similar to a *chatbot* but has additional functionalities such as playing activities, and maintaining a context of interaction, including the history of the past interaction (in order not to repeat itself), physical context (to tailor interaction to the date, time, weather...), and a profile of the user (to personalise interaction to them). Contrarily to a chatbot which is designed for short-term interactions, we expect such a *companion* agent to be able to develop a long-term relationship with the user. This will be tested with a Wizard of Oz setting before our industrial partner provides us with a children-specific ASR.

The dialogue above is obviously not as rich as child-mother interactions from the CHILDES corpus; in particular it lacks the recognition of emotions and expression of empathy that is essential in human interactions. Therefore future directions for research include detecting the child's emotions (we have been experimenting with OpenEar (Eyben et al., 2009) to detect emotions from voice); reasoning about detected emotions, using an existing BDI model of emotions (Adam, 2007); helping the child to cope with them, in particular by showing empathy; and endowing the toy with its own personality (Goldberg, 1993).

6 Acknowledgements

This project is supported by the Australian Research Council, and RealThing Pty Ltd. under Linkage Grant LP0882013

References

F. Abbattista, G. Catucci, M. Degemmis, P. Lops, G. Semeraro, and F. Zambetta. 2003. A framework for the development of personalized agents. In *KES*.

C. Adam. 2007. *Emotions: from psychological theories to logical formalisation and implementation in a BDI agent*. Ph.D. thesis, INP Toulouse, France.

P. B. Brandtzaeg, A. Folstad, and J. Heim. 2006. Enjoyment: Lessons from karasek. In M. A. Blythe, K. Overbeeke, A. F. Monk, and P. C. Wright, editors, *Funology: From Usability to Enjoyment*. Springer.

F. Eyben, M. Wollmer, and B. Schuller. 2009. openEAR: Introducing the Munich open-source emotion and affect recognition toolkit. In *ACII*, Amsterdam.

D. Field, R. Catizone, W. Cheng, A. Dingli, S. Worgan, L. Ye, and Y. Wilks. 2009. The senior companion: a semantic web dialogue system. (demo). In *AAMAS*.

J. Ginzburg. 1997. Resolving questions I and II. *Linguistics and Philosophy*, 17 and 18.

L. R. Goldberg. 1993. The structure of phenotypic personality traits. *American Psychologist*, 48:26–34.

B. MacWhinney. 1995. *The CHILDES Database*.

B. MacWhinney. 2000. *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum Associates.

R. Ronnquist. 2008. The goal oriented teams (gorite) framework. In *Programming Multi-Agent Systems*, volume LNCS 4908, pages 27–41. Springer.

O. Stahl, B. Gamback, M. Turunen, and J. Hakulinen. 2009. A mobile health and fitness companion demonstrator. In *EACL*.