

Managing Domain-Oriented Spoken Conversation

Niels Ole Bernsen

NISLab,
University of Southern Denmark
nob@nis.sdu.dk

Laila Dybkjær

NISLab,
University of Southern Denmark
laila@nis.sdu.dk

Abstract

So far nearly all spoken dialogue systems (SDSs) have been task-oriented. An important question for the future of SDSs is what happens to dialogue management if we move beyond the task-oriented system to domain-oriented systems which are no longer defined in terms of the tasks they support. This paper discusses conversation management for domain-oriented systems by describing the solution adopted in the recently completed first prototype of a system which enables primarily 10-18 years old users to have conversation with life-like animated fairytale author Hans Christian Andersen.

1 Introduction

The term ‘dialogue management’ is standard for the spoken dialogue system (SDS) functionality which interprets the user’s input in the task and emerging discourse contexts, reasons about the input when needed, retrieves necessary information from external information sources, such as databases, and decides on the system’s next output. A dialogue manager component is located at the core of the system, receiving input from an input processing chain consisting of a speech recogniser and a natural language understanding module, and producing output for an output processing chain which includes, at least, a text-to-speech (TTS) system, or speech synthesiser, and sometimes also a response generator which turns the dialogue manager’s output into TTS surface language [Bernsen et al. 1998]. Today, dialogue managers are at work not only in unimodal (speech-only) SDSs used via the telephone or a microphone in stationary environments but also in multimodal SDSs, in mobile and ubiquitous SDSs, and for a multitude of applications, including information systems,

tutorial SDSs, etc. One property common to virtually all these systems, whether commercial or research systems, is that they are *task-oriented*, i.e. they help users solve one or several tasks through spoken and sometimes multimodal interaction.

From a system developer’s point of view, task-orientation is an extremely helpful property. The nature, combinatorics, and inherent logic of the task, and the assumption that user and system share the sole goal of completing the task during dialogue, provide a strongly constraining environment on users’ vocabulary, grammar, and information input, which again makes it possible to develop dialogue managers for even rather complex tasks. Today’s task-oriented dialogue managers are far from having reached the level of standardisation, see, e.g., [Charfuelàn and Bernsen 2003] for a review of current approaches. Some trends are notable, however, such as towards dialogue manager modularity and separation between task-dependent and task-independent processing, both trends supporting portability to new applications, and the trend to separate overall system management and dialogue management, for instance by using a hub architecture.

An important question for the future of SDSs is: what happens to dialogue management if we move beyond the task-oriented system? Given the definition of task-orientation above, what happens first, so to speak, is that all task constraints disappear. Instead, we face development tasks which cannot rely on shared-goal collaborative task resolution among user and system. In order to begin to conceptualise the vast space of, so far, only negatively defined SDSs, i.e. *non-task-oriented* SDSs, we propose two goalposts. The ultimate goalpost is that of Turing test-compliant systems which can have human-style conversation about virtually everything [Turing 1950]. While these systems are still a long way off, the intermediate goalpost of *domain-oriented systems* would seem to constitute a worthwhile state of the art challenge. In a first, tentative, definition, a domain-oriented system is a system which can conduct conversation about almost everything within one or several

particular domains. We use the term ‘conversation’ to refer to the spoken interaction which takes place when there is no common task to be solved among the interlocutors. Presumably, the absence of a common task implies the absence of a single shared user-system goal in any operational sense.

This paper discusses conversation management for domain-oriented systems by describing the solution adopted in the recently completed first prototype (PT1) of a system which enables users to have conversation with life-like animated fairytale author Hans Christian Andersen. Section 2 presents an overview of the system. Section 3 discusses the underlying conception of conversation. Conversation management is described in Section 4. The concluding section (Section 5) discusses open issues of evaluation and otherwise, some of which we hope to learn more about in the upcoming first prototype user test and all of which we would like to address in the second prototype.

2 The NICE HCA system

The Hans Christian Andersen (HCA) system is being developed in the NICE project (2002-2005) on Natural Interactive Communication for Edutainment [NICE]. The project has five partners: LIMSI (Paris, France), Liquid Media (Stockholm, Sweden), NISLab (Odense, Denmark), Scansoft (Aachen, Germany), and TeliaSonera (Stockholm, Sweden). Scansoft delivers speech recognition optimised for children and adolescents for Swedish and English. LIMSI is in charge of gesture recognition and interpretation, and input fusion. Liquid Media is responsible for animation and system integration. NISLab develops natural language understanding, character modelling, and response generation for HCA. TeliaSonera develops natural language understanding, character modelling, and response generation for some of HCA’s fairytale characters. speech synthesis is off-the-shelf software. In PT1, conversation with HCA is in English while spoken interaction with the fairytale char-

acter is in Swedish. In the second prototype (PT2), the user should be able to go from HCA’s study into the fairytale world. In PT1, the two worlds remain separate. In this paper focus is solely on the HCA system.

The architecture of the HCA system is shown in Figure 2.1. The modules communicate via a central Message Broker which is publicly available from KTH [KTHAgentComm]. The Broker is a server that routes function calls, results and error codes between modules. The Transmission Control Protocol (TCP) is used for communication. The Broker coordinates input and output events by time-stamping all messages from the modules as well as associating them to a certain dialogue turn. The behaviour of the Broker is controlled by a set of message-passing rules, specifying how to react when receiving a message of a certain type from one of the modules. In PT1, speech recognition is simulated. The speech recogniser is not yet linked up to the rest of the system and also needs to be trained on children’s voices. Pre-PT1 Wizard of Oz acoustic data as well as data collected with PT1 will be used for training purposes.

The primary use setting of the HCA system is in museums and other public locations at which interactions with an international user audience are expected to have an average duration of, say, 3-15 minutes. Target users are 10-18 years old children and teenagers. The system is an edutainment system. It attains its educational goal by providing correct factual information, both visually and orally. However, an equally important goal is to entertain through emulated human-human conversation. Figure 2.2 shows 55-year old HCA surrounded by artefacts in his study and doing various gestures. The study is a rendering of HCA’s study on display in Copenhagen, modified so that he can walk around freely and so that a pair of doors lead into the fairytale world. Also, pictures relating to HCA’s knowledge domains have been hung on the walls.

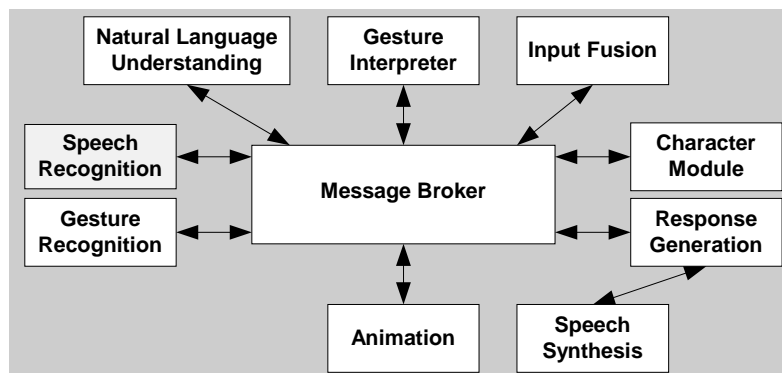


Figure 2.1. General NICE HCA hub system architecture.



Figure 2.2. HCA in his study.

The user can address, in any order, any topic within HCA's knowledge domains, using spontaneous speech and mixed-initiative dialogue. In PT1, the domains are: HCA's works (only his fairytales), his eventful life, his physical presence in his study, the user, and HCA's role as "gate-keeper" for access to the fairytale world. In addition, HCA has a 'meta' domain in order to be able to handle meta-communication during conversation. When the dialogue is about physical presence, the user may also use 2D gesture input to indicate an artefact, i.e. a picture or some other visible object, which HCA might want to tell a story about. HCA reacts emotionally to the user's input, e.g., by getting angry or sad because of what the user says, or by getting happy if the user likes to talk about his fairytales, cf. Section 4.4.

In PT1, HCA has a fairly limited knowledge about his domains of conversation. We have implemented his knowledge domains breadth-first in order to explore the issues involved, rather than implementing a single domain in depth. The cover story is that HCA is coming back! However, he still has to re-learn much of what he once knew. If the user would do him the favour of visiting him later, he is convinced that he will have become much more of what he once was. In addition to the very true information provided by this cover story, the story

may help convince users that HCA is not (yet) a full virtual person and thus make them behave accordingly. HCA does not tell the cover story up front to new users and does not, more generally speaking, instruct users on how to interact with him or inform them of what he is able to engage in conversation about. Rather, users will be told his cover story if they either explicitly ask what HCA knows about or can do, or if they show too much interest in things he does not know about (yet).

3 On the nature of conversation

In the last few years, there has been a tendency, not least in the animated interface agents community, to identify talking agents with conversational agents [Cas-sell et al. 2000]. Understandable as this may be, most talking agents today carry out task-oriented dialogue rather than conversation in the proper sense of this term. When we talk to someone to get information or to coach, we do not have conversation with him or her. Conversation is different, but what is it?

So far, our search for relevant literature on the general nature of conversation has failed to turn up any authoritative account, except for, e.g., healthy practical advice on how to make friends, e.g., [Gabor 2001]. This

suggests that the SDS community may have to lay part of the groundwork itself. On closer inspection, conversation, properly so-called, turns out to be very different from task-oriented dialogue. By comparison with task-oriented dialogue with its shared-goal principles of cooperativity, serving to maximise efficiency and successful task completion [Berssen et al. 1996], spoken conversation has a different, and often contrary or even contradictory, richness. Based on our analysis, HCA has been designed to follow a set of *principles of prototypical successful human-human conversation*. The term ‘prototypical’ means that we are trying to capture, at least, part of the essence of “a good conversation”. The principles are:

1. initially, in a polite and friendly way, the interlocutors *search for common ground*, such as basic personal information, shared interests, shared knowledge, and similarity of character and personality, to be pursued in the conversation;
2. the conversation is successful to the extent that the interlocutors *find enough common ground* to want to continue the conversation;
3. the interlocutors provide, by and large, *symmetrical contributions* to the conversation, for instance by taking turns in acting as experts in different domains of common interest, so that one partner does not end up in the role of passive hearer/spectator, like, for instance, the novice who is being educated or trained by the other(s);
4. to a significant extent, the conversation is characterised by the participants taking turns in *telling stories*, such as descriptions of personal experiences, observations, anecdotes, descriptions of items within their domains of expertise, possibly jokes, etc.;
5. conversation is *rhapsodic*, i.e. highly tolerant to digression, the introduction of new topics before the current topic has been exhausted, etc. Yet conversation also requires a reasonable amount of conversational control and *coherence* in order not to fall apart into disjointed semi-monologues;
6. conversation, when successful, leaves the partners with a sense that it has been *worthwhile*.

It may be noted that the above principles do not mention entertainment at all, despite the fact that the HCA system has an *edutainment* goal. This is partly because we assume that successful conversation is entertaining in itself, and partly because we want to increase the focus on entertainment in the second HCA prototype with assistance from a professional scenographer who is used to work with the 10-18 years old.

The way in which HCA pursues the six principles (P) above is the following. HCA assumes, of course,

that the user takes some interest in his life and fairytales as well as in himself and his study (P1,P2). However, he is aware that common ground is a dual-aspect notion, having both an HCA aspect and a user aspect. As for the user aspect, HCA asks polite questions about the user early on, such as about the user’s name, age, gender, and nationality. He also tries to elicit user opinions on his fairytales, his visible persona, and his study. These HCA initiatives serve the goal of conversational symmetry as well (P3), as does the following. HCA puts an effort into making the user the expert in conversation by asking about games played by children and adolescents today, demonstrating interest in football, computers, and otherwise. During a Wizard of Oz (WoZ) field collection of 30 hours of approx. 500 spoken dialogues with mostly young users at the HCA Museum in Odense in the summer of 2003, we found that the users did show keen interest in telling HCA about contemporary game-playing. They were equally happy telling him about technical inventions made after HCA’s times, seeing that HCA takes a keen interest in photography, trains, and other contemporary inventions. HCA, in his turn, does not just answer questions, or ask them, but tells stories and anecdotes and offers personal views – about his life, his fairytales, about wall pictures and other objects in his room, etc. (P3,P4).

HCA’s main problem in conducting human-style conversation seems to be that he cannot always pursue in depth a topic launched by himself or his interlocutor because, at this stage of development, at least, (i) his knowledge and conversational skills are still somewhat limited, and (ii) we do not have sufficient information about the key interest zones of his target audience. This is where the rhapsodic nature of conversation (P5) may come to his rescue to some extent. When, during conversation, and despite his conversational agenda (Section 4), HCA gets lost and does not understand what the user is saying, he changes topic or even domain in order to try to recover some amount of conversational control. It is in this situation that he also cracks (non-situated) jokes. At this point, however, it is unknown to which extent HCA’s rhapsodic behaviour may catapult him into conversational incoherence and semi-monologue. Judging from the WoZ corpus referred to above, users in the target group are pretty tolerant to digression. However, users who insist on pursuing a topic beyond HCA’s current knowledge get frustrated and tend to thwart HCA’s attempt at re-gaining control.

The upcoming user test of the first HCA system prototype (January 2004) will provide information on the extent to which our implementation of the conversational strategies described above promise to achieve domain-oriented conversation, including evidence on whether the conversation is considered worthwhile by the users (P6).

4 Conversation management

In this section, we describe how the design of the NICE HCA PT1 system aims to meet the requirements to prototypical successful conversation described in Section 3. Figure 4.1 shows the architecture of the system's Character Module which also serves as conversation manager. Some figures to be kept in mind in the following are that HCA in PT1 can produce: spoken output based on some 300 different spoken output templates; some 100 different elementary non-verbal behaviours, such as eye blinks, smiles, pointing gestures, or walk-arounds; and spoken/non-verbal output concerning some 20 objects in his study. Obviously, the elementary non-verbal behaviours can be combined into a virtually unlimited set of complex behaviours, such as highly expressive metaphorical gestures. HCA uses such complex behaviours to visually underpin his stories.

4.1 Non-communicative action and communicative functions

HCA is always on-screen in his study whether or not he is having conversation with a user. The life-likeness design constraint implies that he must be life-like also when not engaged in conversation. We call this system output state *Non-Communicative Action* (NCA). In this state, HCA goes about his day-to-day work in his study, writing, reading, looking out the window onto the streets of Copenhagen, gesturing a bit, etc. In PT1, HCA's non-interactive NCA state has been realised in a rather basic way.

When a user addresses HCA through speech and/or gesture, HCA exits the NCA output state and enters the *Communicative Function* (CF) [Cassell et al. 2000] output state. In this state, he shows attention to the user's input, looking at the user, nodding, gesturing, muttering "yes", "OK" and the like. The CF state is "fast-track" in the sense that it is elicited as soon as the user starts speaking to, and/or gesturing towards, HCA. A problem in the current state of the art is that HCA does not process the user's input incrementally but must show input attention *without* knowing what the user is actually saying or doing. A human, by contrast, would react immediately and often non-verbally to, e.g., user insults or praise. HCA cannot do that and must, therefore, receive any user input with a "poker face and body" which, as far as the expression of emotion is concerned, remains unchanged from what it was during his previous output state. In human terms, however, this behaviour might actually be perceived as respectful conversational behaviour, so we might not need to worry too much about user perception of HCA's CF behaviour. Like HCA's NCA behaviour, his CF output is rather basic in PT1.

4.2 Communicative action

Whilst HCA is showing attention to the user's input, the HCA Character Module is simultaneously processing this input. When the processing is finished, HCA goes into his third output state which we call *Communicative Action* (CA). The output state transitions are managed by the Character Module (CM) Manager which also takes care of module-external communication. It is in the CA state that HCA produces output in response to the user's input. User input processing is done by the Character Module's Mind State Agent (MSA) in consultation with the Knowledge Base, cf. Figure 4.1. MSA processing is controlled by the MSA Manager. The MSA Manager receives a frame from the CM Manager containing the input sent by the Input Fusion module to the CM Manager wrapped in XML. The frame is adapted for internal use in the processing by the MSA modules. The frame incrementally gets filled with various information needed by one or more of the MSA modules and, eventually, with the output references to be sent to the Response Generator.

The general approach to conversation management explored in the NICE HCA PT1 system is the following. Since, due to its rhapsodic nature, conversation can go anywhere (Section 3); since a substantial minimum of conversational coherence is required by each interlocutor (Section 3); and since, by definition, conversation design cannot rely on task constraints (Section 1); the system needs, first of all, a general grasp of the conversation in terms of what has been addressed so far, and what it might be appropriate to address next, including whether or not to respond to the user's input and how to possibly continue the conversation. This general grasp is embodied in the agenda of the Conversation Intention Planner (CIP, Figure 4.1).

At the overall level, the CIP (i) decides whether or not to respond to the user's input and (ii) proposes how to continue the conversation. Apart from looking at the input topic(s) and domain(s) determined by the Natural Language Understanding module and the Gesture Interpreter based on the user input semantics, the CIP performs no processing of the user's input. This is left primarily to the Domain Agents. The information about input topic(s) and domain(s) is used by the CIP to decide and propose how to continue the conversation. The decision is based on HCA's conversation agenda as represented in the CIP. When a new conversation starts, a list of which domains and topics can be used as conversation continuations is uploaded from the knowledge base to the CIP. Throughout a conversation, the CIP keeps track of which of these continuations have been used already in order for HCA not to repeat himself and appear senile. The conversation agenda also includes semi-ordered preferences for the order in which HCA

wishes to address the particular domains, and principles for when to change domain or topic.

On HCA's conversation agenda, the meta domain always takes priority because HCA wants to resolve any meta-communication issues before proceeding. It may be noted that the handling of meta-communication is somewhat different from what one typically finds in task-oriented systems. For instance, HCA replies to insults by also insulting the user, and if he fails to understand the user he may decide to simply talk about something else.

The user domain has high priority because HCA wishes to get to know the user early on in the conversation. The information collected about the user is stored in the User Model. The knowledge about the user's age is used when selecting output in certain cases.

The gatekeeper domain, on the other hand, has low priority because HCA wants to continue conversation on his favourite domains before risking to lose the user to the fairytale world. Thus HCA will try to stick to other domains until they have been reasonably covered in the conversation before he accepts to let the user into the fairytale world. Domain changes are basically only made when a domain has been reasonably addressed already. Topic changes happen when HCA is not pleased with some topic raised by the user, such as his age, or when a topic is exhausted, so that HCA has nothing more to say about it. In addition, HCA has some favourite domains and topics which he will go to some length to pursue in conversation, such as his fairytales.

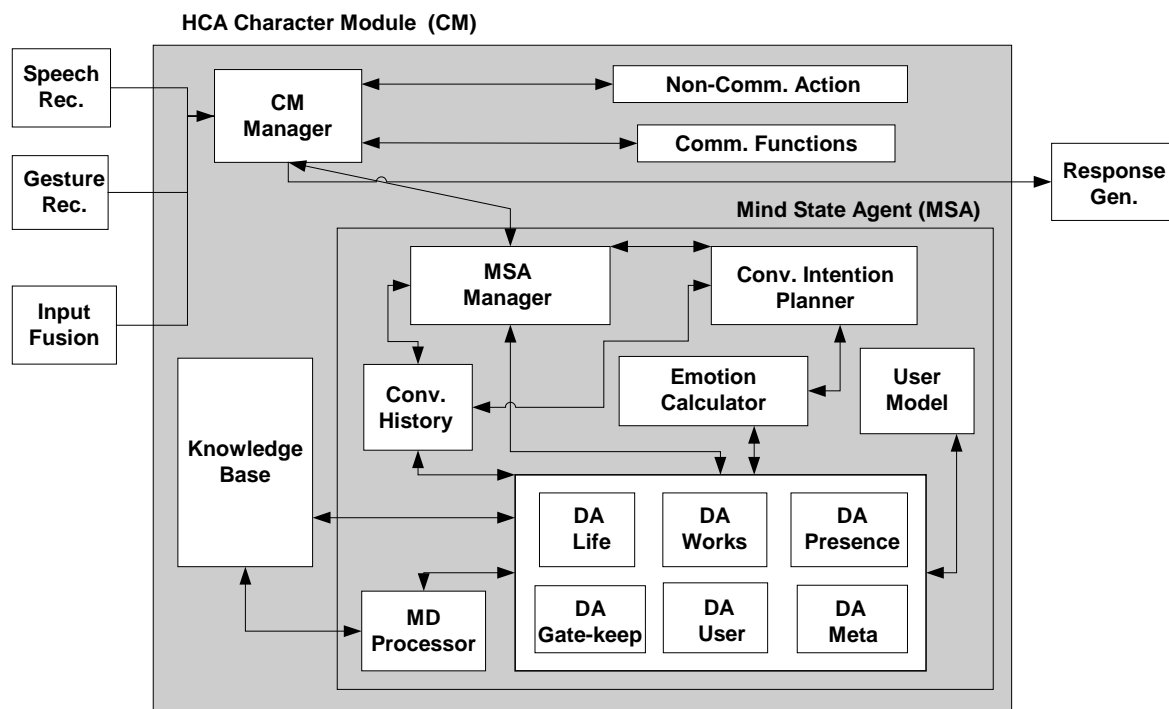


Figure 4.1. HCA character module. DA means domain agent. MD means mini-dialogue.

A CIP-proposed continuation may be a question, a statement, or a mini-dialogue. A question has the function of maintaining HCA's control of the conversation. A statement offers an opportunity for the user to seize initiative. A mini-dialogue is a predefined small dialogue, task-oriented spoken dialogue-style, that will allow HCA, on occasions at which he takes particular interest in the user's input, to carry out in-depth conversation on (a) certain topic(s).

The Mind State Agent Manager (MSAM) receives in a frame the CIP's decision concerning whether to respond to the user's input as well as the CIP's proposed continuation. If no reply to the user's input is to be looked up in the Knowledge Base, the MSAM will al-

ways use the continuation. Otherwise, the MSAM will first try to retrieve a reply from the Domain Agents. If the reply is empty, it will use the continuation. If the reply is non-empty, the MSAM will randomly decide whether or not to use the continuation. To identify a reply and to find the details about a continuation, the MSAM must contact the Domain Agents (DAs). Three functions are available for these purposes:

- function *get reply* takes as input a frame containing the user's input and HCA's emotional state. The return may be an output reference obtained from the Knowledge Base or null in case there is no reply in the Knowledge Base for the user input;

- function *get continuation* takes as input a frame which, among other things, indicates the type of continuation, i.e. whether it is a statement, a question, or a mini-dialogue. The return from the function will always be an output reference retrieved from the Knowledge Base;
- if the continuation concerns meta-communication, a function with a parameter action is used that indicates the particular meta case to be handled, e.g., *repeat* or *insult*. The return from the function will always be an output reference retrieved from the Knowledge Base.

The system has six domain agents (DAs), one for each of HCA's five knowledge domains and one for the handling of meta-communication. The DA component will internally direct the request from the MSAM to the relevant DA. The DAs perform the domain reasoning needed and handle input semantics processing in one of two ways. The user's input is either sent to (i) the Knowledge Base as an SQL query in order to retrieve HCA's reply based on input domain, topic, and semantics, or – only in case of a mini-dialogue – to (ii) the Mini-Dialogue (MD) Processor. The Mini-Dialogue Processor is a finite-state machine which processes the input in the dialogue context and produces an internal identifier which the DA uses to query the Knowledge Base and retrieve a reference to HCA's output.

In addition to the CIP's maintenance of a record of which continuations have been used already, a Conversation History (CH) is accessible to all main Mind State Agent modules. The Conversation History keeps track of information about the individual input and output turns in the conversation, for instance in order to enable HCA to repeat what he just said, the number of consecutive turns involving meta-communication, and the mini-dialogue status (started, ongoing, ended). Keeping track of numbers of turns of type T(n) enables the CIP to take specific action when, e.g., the user has produced several low-confidence score input turns.

The Emotion Calculator (EC) updates HCA's emotional state whenever the user's input produces an emotion increment which makes HCA more happy, sad, or angry, cf. Section 4.4.

4.3 Gesture and input fusion

Output from the Natural Language Understanding and Gesture Interpretation modules reach the character module through the Input Fusion module (Figures 2.1 and 4.1). The first NICE HCA prototype only performs limited gesture and combined speech-gesture input processing. Essentially, the Input Fusion module simply combines an n-best gesture interpretation frame from the Gesture Interpreter, if any, with the one-best frame from the Natural Language Understanding module. The combined frame is sent to the Character Module which

performs basic input fusion, when needed. The n-best gesture interpretation frame includes up to five pairs, each consisting of a single object which the user may have indicated through gesture and the corresponding confidence score that this is the case.

In PT1, input fusion is done by the physical presence Domain Agent (Figure 4.1) which queries the Knowledge Base for any story HCA might have to tell about the top-confidence object. Thus, the user's 2D gesture input provides HCA with yet another opportunity for telling stories, this time about objects in his study. User-indicatable objects range from wall pictures of relevance to HCA's life and fairytales through objects relevant to his travels, to utensils such as his feather pen and the odd chair, table, and lamp.

4.4 Emotions

HCA has the simple emotional state space model. His default emotional state is friendly, which is how he welcomes a new user. During conversation, his emotional state may shift towards happiness, sadness, anger, or a mixture of anger and sadness. HCA's emotional state changes as a function of the user's input semantics. A user who likes to talk about the Ugly Duckling, for instance, produces an emotion increment towards happiness. Emotion increments are attached to the output references stored in the Knowledge Base. Whenever the domain agents query the Knowledge Base, it is checked if there is an emotion increment. If this is the case, the EC is called and returns an updated emotional state. Since, in a number of cases, the exact phrasing of output depends on HCA's current emotional state, the Knowledge Base is then queried again with this new emotional state.

As in humans, the strength of HCA's non-default emotions decreases over time. With each user input which does not elicit any emotion increments, HCA's emotional state goes a nudge towards the default state. HCA expresses his emotional state verbally as well as non-verbally.

5 Conclusion: evaluation issues and next steps

There are many challenges and unknowns involved in developing a conversational domain-oriented system like the one described above. Implementation of PT1 was completed by the end of 2003. The system is now being tested to make sure that it works as robustly as possible before the controlled user test in early 2004. The user test will involve 16 children between 10 and 18 years: eight females and eight males, eight 10-14 year olds and eight 15-18 year olds. Analysis of the user test data is expected to help us improve the system and will, together with ongoing analysis of the much larger

Wizard-of-Oz corpus from summer 2003, form the basis for our design decisions for the second prototype (PT2).

Preliminary results of the WoZ data analysis are emerging. The WoZ was a field simulation in which the wizards simulated the PT1 specification and were allowed to perform specific kinds of improvisations along the way. The data suggests that the simulated PT1 appealed more to the target user group of 10-18 year olds than to younger or older users. The first conversation between simulated HCA and a 14 years old boy, Marius, from Norway, is shown in the Appendix. Marius was so satisfied with this 58-turn conversation with HCA that he came back later and had a second, 99-turn conversation with the fairytale author. The combined 157-turn conversations took approx. 14 minutes.

Intuitively, HCA's first conversation with Marius seems to be quite successful in many respects. Clearly, Marius must have found the conversation edutaining or entertaining or he would not have spoken to HCA for so long. Given the theory of prototypical successful conversation in Section 3, a key research challenge is if it is possible to develop a metrics for measuring the extent to which conversations are successful. Measuring the length and duration of conversations is far from sufficient. All it takes to have a lengthy conversation with HCA is a persistent user. Measuring HCA's conversational coherence (cf. Section 3) would seem relatively straightforward. Evaluating its effects on the user is less so but has to be done.

A related issue illustrated by the Marius conversation concerns the extent to which PT1 is able to do what the wizard does in that conversation, and also the extent to which PT2 can be enabled to do what the wizard does. For instance, like PT1 HCA, the wizard uses replies and continuations. However, in PT1 we only keep track of what has been said regarding continuations whereas we do not keep track of replies to user input. More could probably be done to ensure a good coherence between the reply to the user input and the continuation.

Another important evaluation issue is the target users' interest in the domains and topics covered by PT1 and their potential interest in domains and topics not covered by PT1. Extending HCA's PT1 domain coverage is not an end in itself but making edutaining conversation is. We want to know where lies the strongest edutainment payoff in extending HCA's domains and topics for PT2. For instance, the WoZ studies strongly suggest to give HCA a capability of discussing modern inventions with his young users. It is less clear how much more, and exactly what, he needs to know about his fairytales and his life. The planned user interviews along with the upcoming user test should shed more light on this question. A related issue is already on the agenda for PT2. It is to streamline contents maintenance for heavy-duty contents systems, such as the HCA sys-

tem. This will facilitate portability of the system to new applications.

Entertainment is another major issue to look into. We do not know yet how entertaining users find the system. Again, the planned user interviews will provide input on this issue. It is quite possible that the system's entertainment value could be significantly improved by letting a professional scenographer write the output.

There are, of course, many other issues to consider in addition to those already mentioned, such as elaboration on HCA's non-communicative action and communicative functions, and his emotions. However, firm design decisions will only be made once we have analysed the user test data.

Acknowledgement

The work described in this paper is being supported by the European Commission's Human Language Technologies Programme (HLT) under Grant IST-2001-35293. The support is gratefully acknowledged.

References

- [Bernsen et al. 1996]] Bernsen, N. O., Dybkjær, H. and Dybkjær, L.: Cooperativity in human-machine and human-human spoken dialogue. *Discourse Processes*, Vol. 21, No. 2, 1996, 213-236.
- [Bernsen et al. 1998] Bernsen, N. O., Dybkjær, H. and Dybkjær, L.: *Designing Interactive Speech Systems. From First Ideas to User Testing*. Springer Verlag 1998.
- [Cassell et al. 2000] Cassell, J., Sullivan, J., Prevost, S., and Churchill, E. (Eds.): *Embodied conversational agents*. Cambridge, MS: MIT Press 2000.
- [Charfuelàn and Bernsen 2003] Charfuelán, C. and Bernsen, N. O.: A task and dialogue model independent dialogue manager. In Angelova, G., Bontcheva, K., Mitkov, R., Nicolov, N., and Nikolov, N. (Eds.): *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP-2003)*, Borovets, Bulgaria, 2003. INCOMA, Shoumen, Bulgaria 2003, 91-97.
- [Gabor 2001] Gabor, D.: *How to Start a Conversation and Make Friends*. London: Sheldon Press 2001.
- [Turing 1950] Turing, A.: Computing machinery and intelligence. *Mind* 59, 1950, 433-60. Alfred V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling*, volume 1. Prentice-Hall, Englewood Cliffs, NJ.
- [KTHAgentComm] <http://www.speech.kth.se/broker>
- [NICE] <http://www.niceproject.com/>

Appendix

The following transcribed conversation is one of the 498 English conversations between users and HCA recorded in the field at the HCA Museum in Odense, Denmark, in the summer of 2003. The user is 14 years old Marius from Norway, who, shortly after having completed the conversation below, came back and had an even longer conversation with HCA. We consider the illustrated conversation rather successful and are working on making explicit the criteria behind this intuitive judgment. The ultimate aim is to define a metrics for successful conversation.

1: Marius: hello
2: HCA: hello welcome to my study
3: Marius: hello
4: HCA: hello welcome to my study
5: Marius: okay is that your study
6: HCA: this is my study
7: Marius: okay
8: HCA: this is where i write my fairy tales
9: Marius: which of your fairy tales have you been written on your study
10: HCA: all of them
11: Marius: all o' them can you tell me about one o' them
12: HCA: i can tell you about the Little Mermaid the Ugly Duckling and the Princess and the Pea
13: Marius: i haven't heard much about the Little Mermaid could you please tell me 'bout it
14: HCA: yes the mermaid saves a prince from drowning and falls in love with him to become human and get legs. she pays with her beautiful voice because she is now dumb she cannot tell the prince who she is and he marries another causing the mermaid to die
15: Marius: okay a
16: HCA: do do you like the story
17: Marius: yea i like i've watched a c cartoon about it a ca a cartoon movie
18: HCA: what is a cartoon movie
19: Marius: well they use drawings and they use a lot o' them and then they make this whole movie about it
20: HCA: what is your name
21: Marius: my name is Marius
22: HCA: how old are you
23: Marius: i'm fourteen
24: HCA: when i was your age i was a big fan of the theatre whenever there was a possibility i went to the theatre. it was my favourite place number one yea i also like the theatre
25: Marius: it's fun
26: HCA: are you a boy or a girl
27: Marius: i'm a boy
28: HCA: i guess i was not a typical boy

29: Marius: okay
30: HCA: i didn't like boys' games even at school i didn't take part in their games instead i stayed indoors where do you come from
31: Marius: i come from Norway
32: HCA: what kind of games do you like to play
33: Marius: i like music it's not a game but it's one of my favourite things to do
34: HCA: that must be very special can you tell me more
35: Marius: well i have this keyboard and i can make music with drums bass yea
36: HCA: i am curious to know which games children and teenagers play nowadays do they play any games which did not exist when i was a child
37: Marius: i'm not sure i'm a big f fan of chess when it comes to games
38: HCA: what is chess
39: Marius: chess is a game when you use little f figures and you battle against another player
40: HCA: that must be a very special game
41: Marius: yea
42: HCA: at my age i think the game i like the best is to play games with words or make paper cut-outs for small children or charity
43: Marius: yea do you like music
44: HCA: yes i do like music
45: Marius: what kind of music do you like hello
46: HCA: it sure is a nice weather today
47: Marius: yea it's very hot it's the middle of the summer you know
48: HCA: i like the summer
49: Marius: yea me too
50: HCA: do you like my study
51: Marius: yea i like your study
52: HCA: if it was your study would you be missing anything
53: Marius: what do you do if you went out of ink
54: HCA: i will buy some more
55: Marius: oh okay
56: HCA: bye have a nice day
57: Marius: have a nice day you are cool
58: HCA: thank you