Dialogue Modelling for a Conversational Agent

Peter Wallis[1], Helen Mitchard[2], Jyotsna Das[2], and Damian O'Dea[2]

[1] Agent Oriented Software, Pty. Ltd.
Carlton, Victoria
peter.wallis@agent-software.com.au
[2] Information Technology Division
Defence Science and Technology Organisation
Salisbury, South Australia
<firstName>.<lastName>@dsto.defence.gov.au

Abstract. There is growing agreement that dialogue management is critical to speech enabled applications. This paper describes a novel approach to knowledge acquisition in the natural language processing domain, and shows the use of techniques from *cognitive task analysis* to capture politeness protocols from a "dialogue expert." Acknowledging the importance of intentions in mixed initiative systems, our aim was to use an off-the-shelf Belief, Desire, and Intention (BDI) framework from Agent Oriented Software to provide the planning component, and introduce *plan library cards* as a means of capturing expertise in this context.

1 Introduction

Being able to hold a conversation with a computer has been a dream of AI research from the very beginning when Turing proposed what has become known as the Turing Test. It turned out to be harder than expected, and in this year when HAL was to be on his way to Jupiter, the GUI is still the primary means of interfacing to a computer, and call centres employ people to answer telephones.

Two things have changed in recent years that make dialogue more attractive as a research area. First, with the rise of the call centres, there is more research funding available, not only for speech recognition, but also for the software that decides what to say, and when to say it. Second, the research community now accepts there will be no silver bullet, and that a working AI system will require a concerted effort by a team of people doing sometimes dull things.

The work described here is part of an ongoing project to create a conversational agent using the beliefs, desires, and intentions (BDI) architecture introduced by Rao and Georgeff [1]. BDI systems fall within a long tradition in AI of modelling human decision making by selecting plans from a plan library to match current goals. Populating that library is a key issue, and this paper describes our approach to this task.

For the last ten years the natural language processing (NLP) community has been using *corpus analysis* as its primary data acquisition tool. This approach collects a large body of naturally occurring text, and then uses tools such as statistical models [2, 3] or sequential analysis [4] to infer things about text in general. In this paper we introduce a different approach to knowledge acquisition. We use a technique called Cognitive Task Analysis (CTA) in which a subject matter expert (SME) is *interviewed* to discover their thought processes while performing a task. Similar techniques have been used to populate the rule bases of expert systems [5, 6] and Cognitive Work Analysis has been used to develop software agents for system simulation [7]. Mitchard [8] used Cognitive Task Analysis to create BDI models of human decision making in the air operations domain, and, following on from Mitchard, we use *Applied Cognitive Task Analysis* [9, 10] to elicit knowledge from our dialogue expert.

Our SME's task — let us call her KT — was to take bookings for company cars by telephone. Booking cars is, naturally, of little direct use to the Australian Defence Forces and, like many other tasks, is more conveniently done with a GUI. This particular task should be seen as the pilot study for a more useful embodied conversational agent performing data access on behalf of decision makers.

As far as dialogue is concerned, we find that expressions like "OK," "Yea," "I see," and "Really" not only ground knowledge in the shared space, but can also fulfill the goal of encouraging the other party to say more. This technique is, we claim, key to KT's strategy for being polite.

2 Background - The BDI Architecture

Beliefs, Desires, and Intentions, have long been used as an framework for embedded systems. Bratman's original aim [11] was to describe resource-bounded decision making. Architectures based on his writing provide a way to balance planning and reactive behaviour. It provides a model of making decisions with partial knowledge of the environment, and with insufficient time to make the best decision.

Since it was first introduced, the BDI approach has found a niche in the software agent community. Two common themes in the definition of "agent" are autonomy, which suggests goal driven behaviour, and a separation between the agent entity and its environment. The environment, being outside the control of the agent, provides inputs to which the agent may want to react. BDI is designed to pursue goals while at the same time exploiting opportunities as they arise.

A second reason BDI is closely linked with software agents is that, like SOAR [12], it is a candidate model of human cognition. For many years Air Operations Division at DSTO have been using BDI agents to implement the human element in simulations [13]. Such simulations involve classic software agents with a complex task, and programming them is non-trivial. Domain experts are often brought in to verify the behaviour of agents, and these SMEs tend to find the BDI scripts intuitively clear. Why? One explanation is that the BDI approach explicitly models how we humans think others think. It can be seen as an implementation of the folk psychological view that a rational agent will *do* what it *believes* is in its *interests*. This understanding is so ingrained in us humans that it is often difficult to see why it is interesting, hard, or even useful [14].

Using Dennett's example [15], seeing two children tugging at a toy, we know they both *want* it. We reason about other minds in terms of mental attitudes, and the BDI approach attempts to capture models of decision making at that level of abstraction. When pilots look at a BDI plan in Air Operations simulations, what they see makes intuitive sense because it describes what they would expect another pilot to do. Writing agents in terms of BDI utilises our inbuilt human ability to understand, in a common-sense way, other people's behaviour.

3 Background - Dialogue

Probably the most infamous dialogue system is Weizenbaum's Eliza [16]. This system was implemented using quite a simple procedure; the text is read and inspected for the presence of a keyword. If such a word is found, the sentence is transformed in accordance with a rule associated with the keyword, if it is not, a content-free remark or an earlier transform is retrieved. The text from this retrieval or transform is then printed out as the reply. Since 1966 when Eliza first appeared, there would appear to have been general agreement in the AI community that, although an interesting curiosity, the technique Wiezenbaum used did not bear on the nature of dialogue. Although pattern/action rules could implement a Rogerian psychologist, that role was seen as simply an interesting exception with little relevance to more general skills that would allow a machine to, for instance, book cars.

Much of the work on dialogue since then has concentrated on text generation. This kind of dialogue is often described as goal driven and is known as discourse planning. Consider writing this text. As authors, we have a goal to convince the reader of something and some plans and sub-plans on how to do it. The text planning process can be modelled as a hierarchical set of goals that bottoms out with the production of words on the page. Dialogue, by contrast, involves multiple agents who can interrupt and block each other's goals. It has the added complexity of continual plan failure and re-evaluation — something BDI was explicitly developed to handle. Research on the interactive nature of dialogue includes work on the way the "common ground" is developed between the participants [17], the nature and role of obligation, and what Allan calls "practical dialogue" systems [18]. Research on the latter emphasizes the way people use language to cooperatively solve problems. This is seen as not only practical, but also significantly simpler to achieve than general human conversational competence. The work described here falls squarely in this last camp.

As mentioned above, the primary tool of the NLP community is corpus analysis. In the case of dialogue, a popular approach is for researchers to use *sequential analysis* [4] and mark up transcripts with *dialogue moves* [19, 20], or *rhetorical devices* [21]. This is the set of dialogue moves from a research project for a major Telco:

REQUEST-SERVICE, OFFER-SERVICE, EXPRESS-PROBLEM, ASK-DETAILS, CHECK, ACCEPT-REQUEST, REFUSE-REQUEST, GIVE-DETAILS, CORRECT-

Child's plan #176	
goal:	eat
precondition:	near mum
trigger:	hungry
actions:	
	tell mum "I'm hungry"
	get her to approve
	ask her what I can have
	if I like it, continue
	else post goal "eat chocolate"
	get it
	eat it

Fig. 1. The outline of one of a child's plans for getting food.

INFO, ECHO, ACKNOWLEDGE, PARDON, HOLD, FULFILL, SOCIAL, UNCLASSIFIABLE

Although these types of speech act may seem straight forward, the reliability of the mark-up process still raises questions about the validity of many such tag sets. Better classifications and more effective training and instruction manuals are a hot research topic.

Probably the theory of dialogue structure that comes closest to a BDI approach, is that of dialogue as *dialogue games* [22–24].

Here is an example from Mann [23] introducing dialogue games:

1	*I'm hungry.*
2	Did you do a good job on your geography homework?
3	*Yeah. What's to eat?*
4	Let me read it. What is the capital of Brazil?
5	*Rio de Janeiro.*
6	Think about it.
7	*It's Brasilia. Can I eat now?*
8	I'll let you have something later. What is the capital of Venezuela?
9	*Caracas.*
10	Fine.
11	*So what can I eat?*
12	You want some cereal?
13	*Sure.*
14	O.K.

In this dialogue between a mother and child, the child's desire to eat is only satisfied after mum has checked the homework. At line 1 the child instantiates a plan, something along the lines of that in Figure 1, with the goal of eating. At line 2 Mum has a different goal: to check the child's homework. At line 3 the child tries to stonewall Mum's question, and continues on with her plan. Mum is having none of that, and continues the "check homework" game. At line 7 there is evidence that the child has a plan to wear Mum down — the strategy is that if the child asks often enough, Mum will get sick of saying no. At line 8 Mum explicitly tells her that the wear-Mum-down game is not going to work ("I'll let you have something later") and at line 9 the child has abandoned that plan. With line 10, Mum is indicating that her plan to check homework is finished and the child returns to her plan to get something to eat.

Dialogue *games*, in contrast to dialogue *moves*, are explicitly goal based, longer term, and succeed, fail or are abandoned. Dialogue games are consequently not as explicitly "in" the text, and coding schemes that mark up intentions of the speaker have been found to be unreliable. Rather than looking for games in transcripts, we introduce the idea of explicitly asking a "dialogue expert" about the dialogue games they use. Before looking at the study however, it is informative to consider exactly what it is the study intends to achieve.

3.1 Mixed Initiative and Politeness

Mixed initiative is often seen as the "Holy Grail" in the quest for better dialogue systems. Our primary premise is that a BDI architecture will provide the control structure to enable a mixed initiative dialogue. The concept of a dialogue game describes what the required BDI plans would look like, and ACTA provides the tools to populate the plan library. It is still not clear what *kind* of thing we are looking for however. In human to human conversations, why does initiative shift from one participant to another? When can a participant propose a new goal and when are they obliged to stick with the current one? The hypothesis is that *politeness* is a key motivation in initiative shift in human dialogue. Politeness is not just a matter of saying please and thank you. Brown and Levinson in their seminal work [25] list 30 or so universal strategies for maintaining the "face" of conversants. Interestingly many of these strategies are goal based and so, for instance, if a conversant expresses a desire for X, *positive face* can be expressed by the other person if they also consider X desirable.

The importance of getting politeness right is perhaps demonstrated by the Microsoft Paper Clip. It goes without saying that Mr. Clipit is (was) not popular, but on examination it appears to work quite well as a mechanism for accessing the Microsoft help system. So why the user reaction? One explanation is that it is not playing the social games we expect rational agents to play. On reflection, it appears that the Microsoft Paper Clip is annoying rather than ineffectual.

If user satisfaction is a product of both effectiveness and social skills, it is instructive to consider whether social skills can compensate for poor effectiveness. Evidence from our study suggests this is the case. The car booking scenario can be seen as a slot filling task (in the cases were the caller wanted to book a car — see below) in which the aim of the conversation is to fill in a form with five or so slots: name, destination, time, duration, and contact details. One measure of effectiveness in this context is the proportion of data provided by the caller that makes it into the appropriate slot. KT's error rate can be measured as the number of times the caller provides a piece of data that KT does not pick up, divided by the number of pieces of slot fill data provided. Going through the transcripts, it turns out that she misses 20% of the data callers provide. Keep in mind that *KT* was approached for these experiments because she is recognised as being good at her job, and although user satisfaction was not explicitly measured, there seems little doubt people were happier dealing with *KT* than they would have been working with a machine with a 20% fail rate[1] This has significant consequences for organisations that want to improve user satisfaction with their speech enabled systems.

4 A BDI model of *KT*

We wanted to look at *KT* booking cars over the phone as a pilot study for an intelligent assistant project in the Division. Given time limits, car bookings were not going to give enough samples from our Division alone, and so we approached Electronic Warfare Division for assistance. As a carrot we promised a carton of beer (funded from our own pockets) for the Division that made the most phone calls. The beer becomes important.

A separate recording telephone was installed in *KT*'s office and email sent to both Divisions asking people ring that number to book cars rather than doing it through the existing Outlook calendar. Over two weeks there were 25 calls, 2 of which were taken by a stand-in operator while *KT* was away.

KT was told that the aim of the exercise was to look at politeness and that she would be interviewed after the data was collected to see if we could identify her goals and procedures, and what cues she used to select them. The tapes were transcribed, and this shows a transcript (with names changed) of one of the more successful calls that gives a feeling for the car booking process. When looking at transcripts, bear in mind that a dialogue that seems perfectly natural and comprehensible when spoken, can appear quite awkward when transcribed..

1	*Morning ITD KT speaking*
2	Morning KT, it's PD again
3	*Hello, how are you? [laughter]*
4	Can I book the car for 10 o'clock again please?
5	*Yes, which one was it that you like?*
6	Okay, ZKJ292
7	*292. Um for 10.30?*
8	No, no. 10 till 12
9	*10 till 12. And is it to go to the same place?*
10	Yes, same place. Elex Adelaide
11	*Not a problem, I'll put it in.*
12	Thanks for that KT
13	*Okay, thank you, bye*
14	Bye

[1] The reason *KT* misses data is of course the limitations of human memory and attention when trying to use Outlook and hold a conversation at the same time. Computers of course do not have these limitations.

As usual in AI, the straightforward cases are not interesting; it is the exceptions that require common-sense and where AI systems let us down.

4.1 The Knowledge Elicitation Process

Of the various tools under the ACTA banner, it seemed inappropriate to use the Knowledge Audit probes. Dialogue management skills are primarily skills we humans do not need to think about when we use them, and so it seemed inappropriate to ask *KT* what basically she would think was "obvious." Using the transcripts and preliminary interview data, Das and Wallis used their "naïve" understanding of dialogue to produce a *Task Diagram* overview of the task and to identify the cognitively interesting components of the task. Figure 2 provides the sub-tasks that help frame the car booking dialogue process.

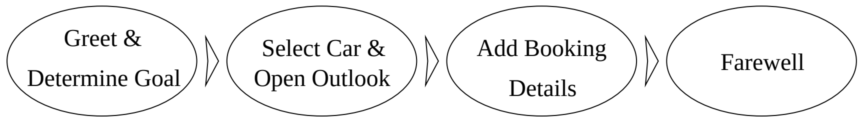

Fig. 2. The Task Diagram for booking-a-car dialogues.

Task diagrams bear a strong resemblance to state transition diagrams, which have been used by some to represent the structure of dialogue for a particular application. Although at this level of description there is a natural order to the sub-tasks, elaborating on the nature of the add-booking-details reveals no such restriction.

The next stage in the analysis was to use techniques from the *Critical Decision Method* (CDM) and ask *KT* why she did things when she did, and to identify her goals when performing some action, her procedures for achieving goals, and the cues she used to initiate procedures and goals. These issues are explored in the context of a "story" and the transcripts provided the context for the interview. In effect the approach was naturalistic observation with supplemental interviews. Phase one was to go through the transcripts and make a first pass at the BDI plans that would implement the necessary dialogue games for the car booking task. Given a set of plans, we could then interview *KT* using probes for CDM to check and develop the model.

4.2 An interesting transcript

Going through the transcripts, the very first call caused problems with identifying the goal. It was from a person who had already booked a car but rang anyway. We suspect the caller was after the beer, but *KT* (being nice) thinks he just wanted to help with the experiment.

Before looking at the transcript, keep in mind that *KT* is expecting callers who want to book a car. Figure 3 shows what happens.

1	*Good afternoon ITD, KT speaking*
2	Oh good afternoon, I have booked a car for tomorrow, a divisional car;
3	*[Right](1)*
4	[I](1) have to ring you here?
5	*Yes*
6	So I booked a COMMS division car, ZKJ292 for 9.30 till 12.00
7	*9.30 to 12.00*
8	We are going to Adelaide
9	*And it was the ZKJ?*
10	Yeah. 292
11	*292. And what was your name?*
12	Ah PD
13	*Right and your extension number?*
14	97313
15	*97313. Um did you want to just wait while I um check that it's available?*
16	I have booked it [not clear] I did this this afternoon before I got the message
17	*[laughter] Okay*
18	Okay
19	*Not a problem*
20	That'll be okay?
21	*Thank you*
22	Okay, thank you KT
23	*Yes, bye*
24	Bye

Fig. 3. Transcript No. 1 — the caller wants the beer.

What is happening in this conversation? Has *KT* not heard the past tense in the callers opening statement? According to our model, what was going on here is that *KT* has no plan that fits with the situation. The initial view was that she was simply going with the plan she had, and getting the details in order to make a booking – a booking she knew, at some level, she was not going to have to make.

There were other cases where the model did not fit neatly with the transcripts, but this paper concentrates on this particular case as it is the most general, and demonstrates how we used CTA in the context of dialogue.

4.3 The interview

The interview threw a new light on the situation. We used probes similar to those in O'Hare et al. [26]. Looking at the transcripts, *KT* was asked things like "What were your specific goals when you said this?" and "What else might you have said at this point?"

When asked what was going on in the transcript in Figure 3, she said that she was thinking "Oh no! what am I going to do here!" She pointed out that she was aware the car was already booked and that indeed she had used the past tense on line 9. There was no intention to get all the details for a car booking, and even when pressed she would not state an actual goal that would fit with the Dialogue Games approach. So what motivated her responses? If she had decided to go with the plan she had, shouldn't she have been able to say as much? One might posit subconscious goals, but that would not be in keeping with using BDI as a model of cognition. It seemed that KT uses BDI for goal based behaviour, but when all else fails, she has a plan — enabled and disabled by the BDI mechanism — that simply fills in and encourages the caller to say more. In the same way as Eliza hands the initiative back to the user, it seems KT's goal, for her first 2 or 3 responses at least, is simply to encourage.

At some point in this dialogue — about line 9 perhaps — she has developed a new plan to add to her plan library. Here is a call, the next day, from some one from ITD who is also after the beer:

1 *Good afternoon ITD KT speaking*
2 G'day, my name's AD, I'm also in ITD, over in
3 *Oh yes*
4 Um, we've just booked a car
5 *Right*
6 And ah we got that e-mail, so, uh can we do that ah [laughter] terrible thing?
7 *[laughter] Um yeah. Can I just go through it with you and just check that you've got it booked okay?*
8 Yep, sure
9 *Is that alright? Um which car were you, did you just book?*
...

Some time between line 9 of the first call, and this call, KT has created a plan to confirm someone's booking if they have already made a booking with Outlook, but ring up anyway.

We conclude that a key mechanism for human dialogue is the ability to hand initiative back to the other person and simply encourage the other person to say more. Eliza's success relied upon exploiting this social protocol to the hilt. In a BDI model of dialogue, one plan — in fact the default plan for when a goal is not identified — should be to encourage the user to say more.

Figure 4 is a caller ringing to cancel a booking with KT's stand-in. At line 7 PP has no idea what to do with the caller and, we propose, is simply encouraging the caller to say more. Similarly at lines 13 and 15. Once again, at line 9, there is a tendency to go with whatever plan is even partly appropriate, but it is not clear how this would generalize. In this case PP is likely to have a plan with a strong link between the cue of a registration number being said, and bringing up the appropriate Outlook entry. There is also a very low cost to doing this, and there is also a tendency for people to want more information. All of these may contribute to PP apparently going with the book-a-car plan when the caller obviously doesn't intend to.

Here is a case where KT cannot recognize the destination, and uses the encourage technique:

1 *Good morning, customer service point, PP speaking*
2 Oh, um I'm ringing for KT actually
3 *Yes, KT is*
4 Car bookings, yeah
5 *Yep, I can take that for you*
6 Okay, fine, I've just had a car out
7 *Yep*
8 A CD car, ZKJ292
9 *One moment, I'll just bring that up. Sorry, the car number was?*
10 Ah ZKJ292
11 *Yep, and your name was?*
12 Ah PD. I'm back from Adelaide now, so the car can be reused, like.
13 *Okay?*
14 Okay
15 *Yep*
16 Okay I didn't need it as long as I thought
17 *Righty oh*
18 Okay, thanks
19 *Thank you for letting us know*
20 Bye-bye

Fig. 4. KT's stand-in using the encourage strategy.

1 *... and where were you going to be going?*
2 Ah the, it's called the UWB facility
3 *UWB*
4 Yeah
5 *Facility*
6 Which is on the RAAF Base.. and also be going to store 2
7 *Okay and do you know where the keys are for the car?*

Imagine a more direct approach — popular in computer interfaces — that "helpfully" suggests the known options:

1 *... and where were you going to be going?*
2 Ah the, it's called the UWB facility
3 *The available options are ...*

There can be no doubt *KT*'s approach is dramatically more polite.

4.4 Knowledge Representation

Having analysed the data, the conclusion from the analysis needs to be written down. This is what Klien refers to as "knowledge representation" and Militello and Hutton [9] recommend using a *cognitive demands table* to sort through and analyse the data. For each situation, the table lists the cues and strategies used by the expert, and the common errors a novice might make. In our case the target BDI architecture requires that cues and strategies be associated with procedures. To this end we introduce *Plan Library Cards* (PLCs) which map directly into BDI plan structures. The use of cards was inspired by experience with CRC[2] cards as used in the software engineering community. Figure 5 shows some of the more obvious PLCs for the car booking task. Each card represents a

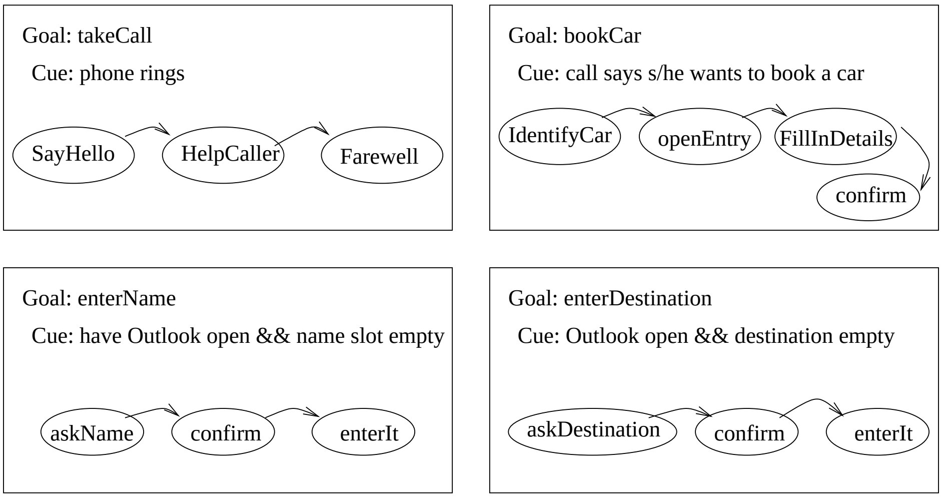

Fig. 5. Four example Plan Library Cards (PLCs) for the car booking task.

procedure; the goal it might achieve; and the cues which determine when it can be used. Note that using the BDI approach, multiple plans might be relevant at any instant but only one is used, and that a procedure can fail or be abandoned at any point — there is no guarantee of completion.

To walk through a transcript, the cards are grouped by goal. When the speaker adopts a goal, the appropriate pile of cards becomes active. Each active pile is then searched for a card with matching cues, and the procedure is executed. That is, in our case, things are said and subsidiary goals are posted. As new cues are discovered, either by looking at the transcript or by interview, they are added to the appropriate card. New cards can be introduced as required and the process repeated until a satisfactory description of the dialogue process is obtained.

Once the analysis is complete the next step is to apply it. Although it would have been nice to implement a phone based car booking system as a demonstrator, we did not have the appropriate resources to do this. We have however been working on the parts of the system that would be portable to other domains. One such component is a Java Speech API [28] based implementation of dialogue which allows "barge-in" statements like those seen in the car booking transcripts. Turning PLCs into an operational system is straight forward using

[2] Class-Responsibility-Collaboration cards. See *UML Distilled* [27, pp64-66] for details.

Agent Oriented Software's product JACK [29] and marrying JACK to the speech system is under way.

5 Conclusion

This paper introduces the use of techniques from Cognitive Task Analysis for knowledge elicitation in the context of BDI systems for dialogue. Intentions are explicitly modelled in a BDI approach, but intentions are hard to capture with more conventional corpus techniques.

We found that one strategy our SME uses is to encourage the other person to say more. It is used when our expert has no plan for furthering shared goals. Such a strategy is more polite than those currently in use in human computer interfaces, and as such would appear to be able to improve user satisfaction independently of system effectiveness.

References

1. Rao, A., Georgeff, M.: BDI agents: from theory to practice. Technical Report TR-56, Australian Artificial Intelligence Institute, Melbourne, Australia (1995)
2. Kupiec, J.: Robust part-of-speech tagging using a hidden markov model. Computer Speech and Language **6** (1992) 225–242
3. Charniak, E.: Statistical Language Learning. MIT Press (1993)
4. Bakeman, R., Gottman, J.M.: Observing Interaction: An Introduction to Sequential Analysis. Cambridge University Press (1997)
5. O'Hara, K., Shadbolt, N.R.: Interpreting generic structures: Expert systems, expertise and context. In Feltovich, P., Ford, K., Hoffman, R., eds.: Expertise in Context. AAAI & MIT Press (1997)
6. Hoffman, R.R., Crandall, B., Shadbolt, N.R.: Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis authors. Human Factors **40** (1998) 254–276
7. Sanderson, P., Naikar, N., Lintern, G., Goss, S.: Use of cognitive work analysis across the system life cycle: Requirements to decommissioning. In: Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society, Houston, TX (1999)
8. Mitchard, H., Winkles, J., Corbett, D.: Development and evaluation of a cognitive model of an air defence operations officer. In Davis, C., van Gelder, T., Wales, R., eds.: Proceedings of the 5th Biennial Australasian Cognitive Science Conference, www.causal.on.net, Causal Productions (2000)
9. Militello, L.G., Hutton, R.J.: Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. Ergonomics **41** (1998) 1618–1641
10. Klein, G., Militello, L.: Cognitive task analysis workshop. Course notes (2000)
11. Bratman, M.E., Israel, D.J., Pollack, M.E.: Plans and resource-bound practical reasoning. Computational Intelligence **4** (1988) 349–355
12. Rosenbloom, P.S., Laird, J.E., Newell, A., eds.: The Soar Papers: Readings on Integrated Intelligence. The MIT Press, Cambridge, Massachussetts (1993)
13. Tidhar, G., Heinze, C., Selvestrel, M.: Flying together: Modelling air mission teams. Applied Intelligence **8** (1998) 195–218

14. Fodor, J.: Psychosemantics. MIT Press, Cambridge, Mass (1987) A Bradford Book.
15. Dennett, D.C.: The Intentional Stance. The MIT Press, Cambridge, Massachussetts (1987)
16. Weizenbaum, J.: ELIZA – a computer program for the study of natural language communication between man and machine. Communications of the ACM **9** (1966)
17. Traum, D.R.: A Computational Theory of Grounding in Natural Language Conversation. PhD thesis, Computer Science, University of Rochester, New York (1994)
18. Allen, J.F., Schubert, L.K., Ferguson, G., Heeman, P., Hwang, C.H., Kato, T., Light, M., Martin, N.G., Miller, B.W., Poesio, M., Traum, D.R.: The TRAINS project: A case study in defining a conversational planning agent. Journal of Experimental and Theoretical AI **7** (1995) 7–48
19. Carletta, J., Isard, A., Isard, S., Kowtko, J.C., Doherty-Sneddon, G., Anderson, A.H.: The reliability of a dialogue structure coding scheme. Computational Linguistics **23** (1997) 13–31
20. Core, M., Ishizaki, M., Moore, J., Nakatani, C., Reithinger, N., Traum, D., Tutiya, S.: Chiba corpus project technical report no. 3. Technical Report CC-TR-99-1, Department of Cognitive and Information Sciences, Faculty of Letters, Chiba University, Japan (1999)
21. Zukerman, I., McConachy, R.: WISHFUL: A discourse planning system that considers a user's inferences. Computational Intelligence **17** (2001)
22. Levin, J.A., Moore, J.A.: Dialogue games: Metacommunication structures for natural language interaction. Cognitive Science **1** (1977) 395–420
23. Mann, W.C.: Dialogue games: Conventions of human interaction. Argumentation **2** (1988) 511–532
24. Levin, L., Ries, K., Thym'e-Gobbel, A., Lavie, A.: Tagging of speech acts and dialogue games in spanish callHome. In: Towards Standards and Tools for Discourse Tagging (Proceedings of the Workshop at ACL'99), Association for Computational Linguistics (1999) 42–47
25. Brown, P., Levinson, S.C.: Politeness: Some Universals in Language Usage. Cambridge University Press (1987)
26. O'Hare, D., Wiggins, M., Williams, A., Wong, W.: Cognitive task analyses for decision centred design and training. Ergonomics **41** (1998) 1698–1718
27. Fowler, M., Scott, K.: UML Distilled: Applying The Standard Object Modeling Language. Addison Wesley Longman, Inc. (1997)
28. Sun-microsystems: The java speech application programmer interface (2001) http://java.sun.com/products/java-media/speech/.
29. Howden, N., Rönnquist, R., Hodgson, A., Lucas, A.: JACK intelligent agents – summary of an agent infrastructure. In: Proceedings of the 5th International Conference on Autonomous Agents, Montreal, Canada (2001)