Hypothesis Generation and Maintenance in the Interpretation of Spoken Utterances

M. Niemann, I. Zukerman, E. Makalic, and S. George

Faculty of Information Technology, Monash University
Clayton, Victoria 3800, Australia
{niemann,ingrid,enesm,sarahg}@infotech.monash.edu.au

Abstract. The DORIS project (Dialogue Oriented Roaming Interactive System) aims to develop a spoken dialogue module for an autonomous robotic agent. This paper examines the techniques used by *Scusi?*, the speech interpretation component of DORIS, to postulate and assess hypotheses regarding the meaning of a spoken utterance. The results of our evaluation are encouraging, yielding good interpretation performance for utterances of different types and lengths.

1 Introduction

The DORIS project (Dialogue Oriented Roaming Interactive System) aims to develop a spoken dialogue module for an autonomous robotic agent. The purpose of this module is to engage in a dialogue with users, and generate plans for the robot that require physical as well as dialogue actions. These plans are created by interfacing with an external planner module. In this paper, we describe *Scusi?*, DORIS's speech interpretation module, focusing on the techniques used to postulate and assess hypotheses regarding the meaning of a spoken utterance.

It is widely accepted that spoken dialogue computer systems are more prone to errors than similar text-based systems. This may be attributed to the inaccuracy of current Automatic Speech Recognition (ASR) systems, disfluencies common in spoken discourse (e.g., false starts, repeats, etc), and the fact that spoken discourse is often less grammatical than written discourse.

In general, during discourse interpretation, a system must be able to (1) postulate promising interpretations, (2) adjust these interpretations dynamically as new information becomes available, and (3) recover from erroneous interpretations. The adjustment and recovery activities are particularly important for the interpretation of spoken discourse. Further, dialogue systems can interact with users to obtain additional information when so warranted by the state of the interpretation process and the dialogue conditions. For example, imagine a situation where an ASR mishears the last part of the request "get me a mug", producing texts that differ mainly in their last portion. If only a few texts are generated, the dialogue module could ask a specific question, e.g., "did you say mug or mud?", but if there are several options, an open-ended question such as "what did you want me to get?" would be more suitable. Similarly, if there are two possible mugs, a reasonable clarification question would be "the green one or the red one?", while if there are several options, then a more appropriate question would be "which mug should I get?". A robot-mounted dialogue system has yet another option — it can direct the robot to perform a physical action (getting closer to a referenced object) instead of asking a question.

The *Scusi?* dialogue interpretation module is designed to support these capabilities. It implements an interpretation process comprising three stages: speech recognition, parsing and semantic interpretation. Each stage in this process produces multiple candidate options, which are ranked according to the probability of them matching the speaker's intention (Section 3). This probabilistic framework supports the (re-)ranking of interpretations as new information becomes available, as well as the identification of 'trusted' (high probability) and 'untrusted' (low probability) regions in an interpretation. The generation and maintenance of multiple interpretations at each stage of the process allows *Scusi?* to recover from erroneous interpretations, and to abstract features of the state of the dialogue which support the generation of appropriate dialogue or physical actions. Features of interest (for each stage of the interpretation process) include: the number of highly ranked interpretations, the similarity between these interpretations, and the probability of each interpretation.

This paper is organized as follows. Section 2 presents an overview of the interpretation process. The estimation of the probability of an interpretation is presented in Section 3, and the process for generating a final interpretation is described in Section 4. Section 5 details the evaluation of the interpretation mechanism. Related research and concluding remarks are given in Sections 6 and 7 respectively.

2 Multi-stage Processing

Scusi? processes spoken input in three stages: speech recognition, parsing and semantic interpretation. In the first stage, the dialogue module runs the ASR software (Microsoft Speech SDK 5.1) to generate candidate sequences of words (i.e., plain text) from a speech signal. Each sentence is assigned a score that reflects the probability of the words given the speech wave. The second stage applies Charniak's probabilistic parser (`ftp://ftp.cs.brown.edu/pub/nlparser/`) to generate parse trees from the plain text sequences. The parser generates n possible parse trees for each text, where each parse tree is associated with a probability. During the semantic interpretation phase, parse trees are successively mapped into two representations based on Conceptual Graphs (CGs) [1]: first *Uninstantiated Concept Graphs (UCGs)*, and then *Instantiated Concept Graphs (ICGs)*. UCGs are obtained from parse trees deterministically — one parse tree generates one UCG (but a UCG can have several parent parse trees). A UCG represents syntactic information, where the concepts correspond to the words in the parent parse tree, and the relations between the concepts are directly derived from syntactic information in the parse tree and prepositions. This process is similar to the assignment of semantic role labels [2].

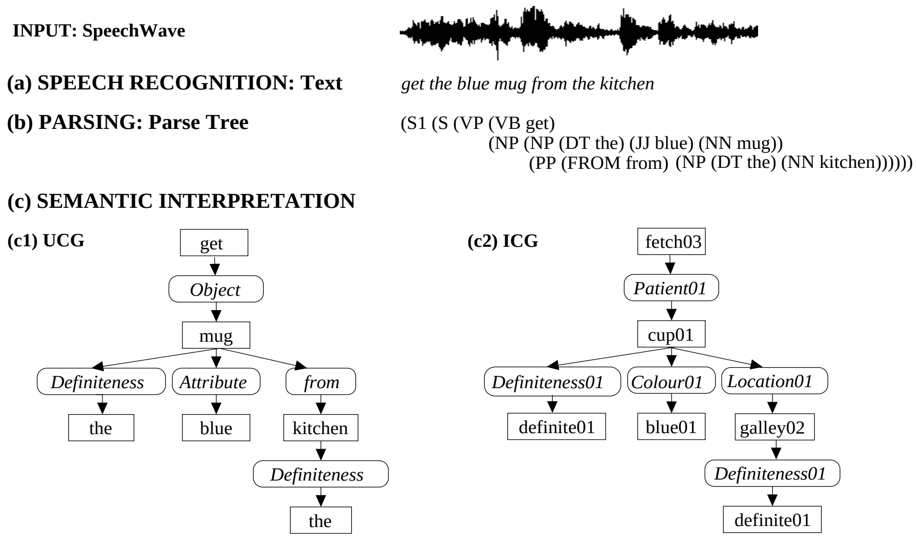

Fig. 1. Structures for the interpretation stages

A UCG can generate several ICGs (and an ICG can have several parent UCGs). An ICG is built by nominating an *Instantiated Concept* from DORIS's knowledge base as a potential realization for each *Uninstantiated Concept* in a UCG (Section 4). Figure 1 illustrates this process with respect to the input sentence "get the blue mug from the kitchen". For instance, the noun 'mug' in the parse tree is mapped to the word `mug` in the UCG, which in turn is mapped to the instantiated concept `cup01` in the ICG.

The consideration of all possible options at each stage of the interpretation process is computationally intractable. Hence, *Scusi?* uses an *anytime* algorithm [3] which applies a selection-expansion cycle to build a search graph. This process is repeated until one of the following is true: all options are fully expanded, a time limit is reached, or the system runs out of memory. At any point after completing an expansion, *Scusi?* can return a list of ranked interpretations with their parent sub-interpretations.

3 Estimating the Probability of an Interpretation

Scusi? ranks candidate ICGs according to their probability of matching the intended meaning of a spoken utterance. The probability of an ICG is estimated from the probabilities of its parent UCGs and the context. Currently, the context encompasses domain knowledge and salience from dialogue history. In the future, the context will also include data from the robot's vision system.

Formally, we denote the user's speech wave by $W \in \mathcal{W}$, where $\mathcal{W}$ is a countable set of spoken utterances. The task is to estimate $\Pr(I|W,\mathcal{C})$, the probability of an ICG I given the speech signal W and the context $\mathcal{C}$. The probability function $\Pr(I|W,\mathcal{C})$ is estimated as

$$\Pr(I|W,\mathcal{C}) \propto \sum_{\Gamma=\{P,U\}} \Pr(I|U,\mathcal{C}) \cdot \Pr(U|P) \cdot \Pr(P|T) \cdot \Pr(T|W) \qquad (1)$$

where U, P and T denote UCG, parse tree and plain text interpretations respectively. The summation is taken over all possible paths $\Gamma = \{P,U\}$ from the parse tree to the ICG. This is necessary since a UCG and an ICG can have more than one parent. Charniak's parser and the ASR return an estimate of $\Pr(P|T)$ and $\Pr(T|W)$ respectively. In addition, it is known that $\Pr(U|P) = 1$, since the process of converting a parse tree into a UCG is deterministic. Hence, *Scusi?* only has to estimate $\Pr(I|U,\mathcal{C})$.

Consider an ICG I containing concepts $n^{ICG} \in \Omega_n$ and relations $r^{ICG} \in \Omega_r$ (Ω_n and Ω_r are the concepts and relations in the domain knowledge respectively). A parent UCG is denoted by U and comprises concepts $n^{UCG} \in \Gamma_n$ and relations $r^{UCG} \in \Gamma_r$ (Γ_n and Γ_r are the words and relations from which UCGs are built). The probability of I given U and context $\mathcal{C}$ can then be stated as follows.

$$\begin{aligned}\Pr(I|U,\mathcal{C}) &= \prod_{n^{ICG}\in I, r^{ICG}\in I} \Pr\left(n^{ICG}, r^{ICG}|n^{UCG}, r^{UCG}, \Omega_n^-, \Omega_r^-, \mathcal{C}\right)\\ &= \prod_{n^{ICG}\in I, r^{ICG}\in I} \left\{\begin{array}{l}\Pr\left(r^{ICG}|n^{ICG}, n^{UCG}, r^{UCG}, \Omega_n^-, \Omega_r^-, \mathcal{C}\right) \times\\ \Pr\left(n^{ICG}|n^{UCG}, r^{UCG}, \Omega_n^-, \Omega_r^-, \mathcal{C}\right)\end{array}\right\} \qquad (2)\end{aligned}$$

where the superscript in Ω_n^- and Ω_r^- denotes the sets Ω_n and Ω_r without the concept n^{ICG} and relation r^{ICG} respectively. It is difficult to estimate Equation 2, as each concept and relation in an ICG depends on the other ICG concepts and relations. We therefore make the following simplifying assumptions: (1) the probability of an ICG relation depends only on the corresponding UCG relation, the neighbouring ICG concepts, and the context; and (2) the probability of an ICG concept depends only on the corresponding UCG concept, the neighbouring ICG relations, and the context. These assumptions yield

$$\Pr(I|U,\mathcal{C}) \approx \prod_{r^{ICG}\in I} \Pr\left(r^{ICG}|r^{UCG}, n_p^{ICG}, n_c^{ICG}, \mathcal{C}\right) \prod_{n^{ICG}\in I} \Pr\left(n^{ICG}|n^{UCG}, r_p^{ICG}, r_c^{ICG}, \mathcal{C}\right) \qquad (3)$$

where n^{UCG} and r^{UCG} denote the UCG concept and relation corresponding to the ICG concept n^{ICG} and relation r^{ICG} respectively. We consider the neighbouring concepts of relation $r^{ICG} \in I$ to be its parent concept $n_p^{ICG} \in I \bigcap \Omega_n$ and its child concept $n_c^{ICG} \in I \bigcap \Omega_n$, and the neighbouring relations of concept $n^{ICG} \in I$ to be its parent relation $r_p^{ICG} \in I \bigcap \Omega_r$ and its child relation $r_c^{ICG} \in I \bigcap \Omega_r$.

After applying Bayes rule, and making additional simplifying assumptions about conditional dependencies, we obtain

$$\begin{aligned}\Pr(I|U,\mathcal{C}) \approx &\prod_{r^{ICG}\in I} \Pr\left(r^{UCG}|r^{ICG}\right) \Pr\left(r^{ICG}|n_p^{ICG}, n_c^{ICG}, \mathcal{C}\right) \Pr\left(n_p^{ICG}|\mathcal{C}\right) \Pr\left(n_c^{ICG}|\mathcal{C}\right) \times\\ &\prod_{n^{ICG}\in I} \Pr\left(n^{UCG}|n^{ICG}\right) \Pr\left(n^{ICG}|r_p^{ICG}, r_c^{ICG}, \mathcal{C}\right) \Pr\left(r_p^{ICG}|\mathcal{C}\right) \Pr\left(r_c^{ICG}|\mathcal{C}\right) \qquad (4)\end{aligned}$$

The last two factors in the first (second) row of Equation 4 represent the prior probabilities of the concepts (relations) in an ICG in light of the context. Hence, moving the concept (relation) factors to the second (first) row will not change the equation. We also collapse the two concept (relation) factors into a single one in order to avoid double counting a node that is both a parent and a child (which is a result of the simplification of the structure of an ICG performed for the above calculations). This yields the following equation.

$$\begin{aligned}\Pr(I|U,\mathcal{C}) \approx & \prod_{r^{ICG}\in I} \Pr\left(r^{UCG}|r^{ICG}\right) \Pr\left(r^{ICG}|n_p^{ICG}, n_c^{ICG}, \mathcal{C}\right) \Pr\left(r^{ICG}|\mathcal{C}\right) \times \\ & \prod_{n^{ICG}\in I} \Pr\left(n^{UCG}|n^{ICG}\right) \Pr\left(n^{ICG}|r_p^{ICG}, r_c^{ICG}, \mathcal{C}\right) \Pr\left(n^{ICG}|\mathcal{C}\right)\end{aligned} \tag{5}$$

Ideally, the probabilities in Equation 5 should be estimated from data, which would require the development of a large database of target ICG models corresponding to different speech signals. Since such a database is currently unavailable, we employ a heuristic approach to estimate the necessary probabilities.

Scusi? associates with each ICG I a confidence score which represents the probability of the ICG matching its parent UCG U and the context. This score is a heuristic approximation of Equation 5, and is used by *Scusi?* to rank competing ICGs. Throughout the ICG-generation process, *Scusi?* uses heuristics to boost or reduce the confidence score (Section 4). Three types of penalties (or rewards) can be applied to this score: small, medium and large. Clearly, the effect of applying a reward is to boost the confidence score, whereas a penalty reduces the score. The actual values for the penalties and rewards are empirically estimated.

4 Generating ICGs

The process of generating ICGs from a UCG and estimating their probability is carried out by Algorithm 1, which is composed of two main stages: *concept and relation postulation* (Steps 2–12), and *ICG construction* (Steps 13–19).

Concept and relation postulation. The algorithm proposes suitable candidate ICG concepts and relations from the knowledge base for each UCG concept and relation respectively (Step 3). This yields a list of ICG concepts L_{c_u} for each UCG concept c_u, and a list of ICG relations L_{r_u} for each UCG relation r_u. The algorithm then estimates the probability of the third and first factor of both rows in Equation 5 for each candidate concept and relation. The third factor represents the prior probability of an ICG concept or relation, and the first factor represents the match between an ICG concept or relation and its UCG counterpart.

The third factor is estimated in Step 6 of Algorithm 1. $\Pr(n^{ICG}|\mathcal{C})$, the prior probability of an ICG concept, is estimated using a function of its salience score. This score is obtained from dialogue history (concepts that were recently mentioned are salient). Relations are not influenced by the dialogue history, hence

Algorithm 1 Generating candidate ICGs given a UCG

Require: UCG U comprising unknown concepts c_u and relations r_u, $k_{\max} = 400$
1: Initialize $\mathcal{I}$ {buffer to hold $k_{\max}$ ICGs}
2: **for all** concepts c_u (relations r_u) in the UCG U **do**
3:    Initialize L_{c_u} (L_{r_u}) {list of suitable candidate domain concepts (relations)}
4: **end for**
5: **for all** concepts $c_k \in L_{c_u}$ (relations $r_k \in L_{r_u}$) **do**
6:    Initialize the probability of c_k to $\mathcal{F}$(salience score from dialogue history)
7:    **for all** comparison functions f associated with concept c_k (relation r_k) **do**
8:       Use f to compare c_u to c_k (r_u to r_k) $\rightarrow$ reward (or penalty) for c_k (r_k)
9:    **end for**
10:    Add concept c_k (relation r_k) to the list L_{c_u} (L_{r_u})
11: **end for**
12: Sort list L_{c_u} (L_{r_u}) in descending order of probability
13: **for** $i = 1$ to $k_{\max}$ **do**
14:    Generate the next best ICG I_j by going down each list L_{c_u} and L_{r_u}
15:    Add I_j to the buffer $\mathcal{I}$
16:    Estimate $\Pr(I_j)$ {product of probabilities of all the concepts and relations}
17:    Check internal consistency of I_j {the probability $\Pr(I_j)$ may change}
18:    Sort buffer $\mathcal{I}$ in descending order of the probability of ICG
19: **end for**

at present their prior probability is set to 1. In the future, their priors will be obtained from a corpus.

The first factor is estimated in Steps 7 and 8. $\Pr(n^{UCG}|n^{ICG})$ (and $\Pr(r^{UCG}|r^{ICG})$), the probability of the match between each concept (and relation) in the UCG and a candidate ICG concept (relation), is estimated by means of comparison functions that give penalties or rewards depending on the result of the comparison. At present, the comparisons are based on four attributes: cg-role (relation or concept), called (a list of domain names for a UCG concept or relation), part-of-speech, and relation (e.g., subject, object). For instance, given the UCG concept *cup*, the ICG concepts `mug01, ..., mug05` and `cup01, ..., cup04` match the UCG concept for the first three attributes (concepts have no relation attribute).

ICG construction. The algorithm generates candidate ICGs by selecting one candidate concept and relation from each of the lists L_{c_u} and L_{r_u} in descending order of probability (Step 14). For example, the ICG generated first is composed of the top (highest probability) element of each list, the next ICG contains the next element in one list and the top elements of the other lists, and so on.

Step 16 calculates the initial probability of an ICG: the product of the first factor in Equation 5 (the match between a UCG concept or relation and a candidate ICG concept or relation) and the third factor (the prior probability of concepts and relations). Step 17 estimates the second factor in Equation 5, which reflects the reasonableness of the relationships between neighbouring concepts and relations. This is done by comparing concepts and relations in the ICG to templates in the knowledge base. For example, the knowledge base contains a

Table 1. Example sentences used during system testing

Sentence	Length	Type	#gold
"get the bag"	short	imperative	2
"the book is on the table"	short	declarative	6
"open the door to the cupboard in the lounge"	long	imperative	1
"the large bag in the bathroom is white"	long	declarative	1
"Michael's box of books is on the table in the bedroom"	long	declarative	3

template entry for a possessive relationship between two concepts. Hence, ICGs that contain a possessive relationship between, say, an action concept and an object concept are penalized.

Algorithm 1 generates a buffer containing k_{max} ICGs the first time a UCG is expanded. It then successively returns the next ranked ICG candidate every time a new ICG is requested for that UCG.

5 Evaluation

A proper evaluation of the *Scusi?* dialogue module is a nontrivial task. Initially, the goal was to compare the performance of *Scusi?* to that of other dialogue systems. However, this was soon found to be infeasible as such systems are commonly too domain specific and difficult to obtain. Consequently, we have decided to evaluate the stand-alone interpretation performance of *Scusi?* for different utterance types.

The evaluation test set comprised 39 utterances: 17 declarative (e.g., "the book is on the table") and 22 imperative (e.g., "open the door"). The two groups of utterances were further sub-divided into short (3 words on average) and long (9 words on average). Each utterance was designed to test different aspects of *Scusi?*'s interpretation process. For example, some utterances refer to the same concepts with alternative words (e.g., "wash" and "clean"). Utterances referring to multiple possible referents were also considered. An example of this is the phrase "a book", since the knowledge base contains several books. Table 1 illustrates sample sentences used during testing. For each utterance in the test set, *Scusi?* was allowed to generate a maximum of 300 sub-interpretations (texts, parse trees, UCGs and ICGs).

An interpretation was deemed successful if it correctly represented the user's intention within the limitations of *Scusi?*'s knowledge base. The accuracy of all interpretations was judged manually by inspection on a case-by-case basis, and the successful ICG(s) were regarded as the gold standard (e.g., the ICGs referring to the different books in the knowledge base are considered gold standards for the utterance "get a book"). The last column in Table 1 shows the number of gold standard ICGs according to *Scusi?*'s knowledge base of 144 items: 24 relations and 120 concepts. In our experiments, we recorded whether the gold standard was found, and if so, the rank of the gold ICG on the basis of its probability.

Table 2. *Scusi?* ranking performance for the full ASR and perfect ASR experiments

	Top	Top 3	Top 10	Not found	Average rank
Full ASR	26	32	36	3	1.8
Perfect ASR	31	36	38	1	1.7

To measure the impact of ASR error on *Scusi?*'s performance, we conducted two sets of experiments. The first round of experiments involved the original ASR output data as returned by Microsoft Speech SDK 5.1. All tests were then repeated assuming a perfect ASR system; i.e., only the ASR output that corresponded to the gold standard text was used.

A summary of our results is given in Table 2. For the first experiment set, the gold ICG had the top or equal top rank in 26 out of 39 utterances. In 10 of the remaining 13 cases, the gold standard was located in the top 10 interpretations. About half of these cases were due to ASR and parsing ambiguity. Only in 3 cases *Scusi?* did not find the gold standard ICG. Near perfect interpretation performance was achieved for short declarative utterances, but the performance of *Scusi?* was otherwise invariant to the length and type of examined utterances. It is worth noting that in 6 cases, the ranking of the gold interpretation increased between the ASR output and the ICG, i.e., a lower ranked text produced a higher-ranked ICG. This demonstrates the effectiveness of maintaining multiple options, instead of selecting only the top-ranked alternative.

For the set of experiments that assume a perfect ASR system, the gold ICG had the top rank in 31 cases, and *Scusi?* did not find the gold ICG for one utterance only. In the remaining test cases, the gold ICG was ranked among the top 10 best hypotheses.

In terms of computational complexity, *Scusi?* required 18 expansion cycles on average before the gold ICG was generated in the first experiment. Furthermore, the mean number of ICGs generated prior to the gold ICG was approximately 3.6. This result demonstrates the effectiveness of our search algorithm.

6 Related Work

This research builds on the work described in [4], whose main focus was the consideration of multiple alternatives during the generation of interpretations. The main focus of this paper is the association of probabilities with interpretations of spoken utterances. In particular, we considered how *Scusi?* creates and ranks candidate ICGs given a UCG.

Many researchers have investigated probabilistic approaches to the interpretation of utterances in discourse systems. Young [5] introduced a probabilistic framework for dialogue systems based on Markov Decision Processes (MDPs) and reinforcement learning. However, the system requires user simulation models for the practical application of the technique. Miller et al. [6] consider a three stage system that maps utterances into meaning representation frames, where each stage is modelled as a statistical process. A similar probabilistic approach is examined in [7] and [8], but Pfleger et al. [8] use an empirical scoring function to interpret multi-modal input, rather than a fully probabilistic model. All these systems employ a semantic grammar for parsing. In contrast, *Scusi?* uses generic, syntactic tools, and incorporates semantic and domain-related information only in the final stage of the interpretation process. Knight et al. [9] compare the performance of a grammar-based dialogue system to that of a system based on a statistical language model and a robust phrase-spotting parser. The latter performs better for relatively unconstrained utterances by users unfamiliar with the system. The probabilistic approach and intended users of our system are in line with this finding.

Concept graphs have been employed for discourse interpretation in [10, 11], but *Scusi?* differs from these systems in its use of UCGs as an intermediate stage that is independent from semantic and domain knowledge. From a processing point of view, Shankaranarayanan and Cyre [11] retain the first parse tree that supports an acceptable interpretation, rather than multiple parse trees. Although Sowa and Way [10] allow for multiple interpretations, their system prunes parses that fail semantic expectations. In contrast, *Scusi?* does not apply a filtering mechanism, and allows flawed candidates to undergo a deeper examination.

The current research closely resembles that in [12, 13] in its integration of context-based expectations with alternatives obtained from spoken utterances. Horvitz and Paek [12] focus on higher level informational goals than those considered by *Scusi?*, using a single output produced by a parser as linguistic evidence for their discourse system. Unlike *Scusi?*, Gorniak and Roy [13] restrict the expected input, factoring in domain knowledge at the outset of the interpretation process. This allows their system to process expected utterances efficiently, but makes it difficult to interpret unexpected utterances.

7 Conclusion

We have described *Scusi?*, the speech interpretation module used in the DORIS project. *Scusi?* is a multi-stage interpretation system that maintains multiple options at each stage of the process, and ranks all (partial) interpretations based on estimates of their posterior probability. This work examined the techniques used by *Scusi?* to postulate and assess hypotheses regarding the meaning of a spoken utterance.

Two separate sets of experiments were conducted using original ASR output data, as well as data generated by a hypothetical perfect ASR. Our empirical evaluation shows that *Scusi?* performs well for declarative and imperative utterances of varying length, but ASR performance has a significant impact. The gold ICG model was ranked in the top three best sub-interpretations in most test cases. In addition, the performance of *Scusi?* was largely invariant to the length and type of utterances considered. In the near future, we will expand our data set, focusing on longer and more complex utterances, and investigate the interactions between the components of the system in order to further improve performance.

Acknowledgments. This research was supported in part by the ARC Centre for Perceptive and Intelligent Machines in Complex Environments. The authors also thank Eugene Charniak for his modifications to his probabilistic parser.

References

1. Sowa, J.F.: Conceptual Structures: Information Processing in Mind and Machine. Systems Programming Series. Addison-Wesley, Reading, MA (1984)
2. Gildea, D., Jurafsky, D.: Automatic labeling of semantic roles. Computational Linguistics **28**(3) (2002) 245–288
3. Niemann, M., George, S., Zukerman, I.: Towards a probabilistic, multi-layered spoken language interpretation system. In: Proceedings of the fourth workshop on knowledge and reasoning in practical dialogue systems (IJCAI'05), Edinburgh, Scotland (2005) 8–15
4. George, S., Niemann, M., Zukerman, I., Marom, Y.: Considering multiple options when interpreting spoken utterances. In: Proceedings of the Fifth IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems, Hyderabad, India (2007) 7–14
5. Young, S.J.: Probabilistic methods in spoken-dialogue systems. Philosophical transactions of the Royal Society A **358**(1769) (2000) 1389–1402
6. Miller, S., Stallard, D., Bobrow, R., Schwartz, R.: A fully statistical approach to natural language interfaces. In: Proceedings of the 34th Conference of the Association for Computational Linguistics (ACL96), Santa Cruz, California (1996) 55–61
7. He, Y., Young, S.: A data-driven spoken language understanding system. In: IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU '03), St. Thomas, US Virgin Islands (2003) 583–588
8. Pfleger, N., Engel, R., Alexandersson, J.: Robust multimodal discourse processing. In: Proceedings of the 7th Workshop on the Semantics and Pragmatics of Dialogue, Saarbrücken, Germany (2003) 107–114
9. Knight, S., Gorrell, G., Rayner, M., Milward, D., Koeling, R., Lewin, I.: Comparing grammar-based and robust approaches to speech understanding: a case study. In: 7th European Conference on Speech Communication and Technology (EUROSPEECH 2001), Aalborg, Denmark (2001) 1779–1782
10. Sowa, J.F., Way, E.C.: Implementing a semantic interpreter using conceptual graphs. IBM Journal of Research and Development **30**(1) (1986) 57–69
11. Shankaranarayanan, S., Cyre, W.R.: Identification of coreferences with conceptual graphs. In: Proceedings of the Second International Conference on Conceptual Structures (ICCS '94). Lecture Notes in Computer Science, College Park, Maryland, US (1994) 45–60
12. Horvitz, E., Paek, T.: A computational architecture for conversation. In: Proceedings of the seventh international conference on User modeling (UM'99), Banff, Canada (1999) 201–210
13. Gorniak, P., Roy, D.: Probabilistic grounding of situated speech using plan recognition and reference resolution. In: Proceedings of the 7th international conference on Multimodal interfaces, Torento, Italy (2005) 138–143