

Probabilistic Dialogue Modelling

Oliver Lemon
CSLI
Stanford University
lemon@csl.stanford.edu,

Prashant Parikh
IRCS
University of Pennsylvania
pjparikh@aol.com

Stanley Peters
CSLI
Stanford University
peters@csl.stanford.edu

Abstract

We show how Bayesian networks and related probabilistic methods provide an efficient way of capturing the complex balancing of different factors that determine interpretation and generation in dialogue. As a case study, we show how a probabilistic approach can be used to model anaphora resolution in dialogue¹.

1 Introduction

The use of probabilistic and decision-theoretic inference in dialogue modelling and management has been explored in preliminary fashion by (Pulman, 1996) and (Keizer, 2001). Probabilistic methods look promising when modelling systems where there is uncertainty, and simple true/false judgments obscure some of the subtleties of representation and processing that are required of an accurate model. Dialogue systems are of this nature because uncertainty is present due to speech recognition noise, speech-act uncertainty, and so on. Epistemic uncertainty is rife in dialogue, and probability distributions provide a natural model of the ambiguities that thus arise. For these reasons it is natural to explore probabilistic representations and algorithms in dialogue management, rather than purely deterministic models. We have experience building deterministic dialogue managers (see e.g. (Lemon et al.,

2001; Lemon et al., 2002)) which use deterministic context update rules.

This paper briefly describes our construction of a Bayes Net modelling dialogue context. We will consider a series of examples of increasing complexity involving anaphoric resolution in Section 3.1. We will point out how they are to be resolved intuitively, and then discuss how our Bayesian net fares. We will see that many of the best insights of deterministic approaches (e.g. in the axiomatic BDI tradition and in the planning literature) can be preserved, often in less brittle forms, in a probabilistic setting.

1.1 Probabilistic modelling ideas

Our approach to resolving anaphora (and dialogue moves) was to generate a probability distribution of the random variable of interest (e.g. salience of referent) and then choose the value of the variable corresponding to the highest probability as the interpretation (e.g. the referent). This decision has a theoretical justification that can be found in a theorem in (Parikh, 1990) in the context of his game-theoretic model of language use. The theorem states that under certain conditions (which hold in our context) the correct interpretation of an utterance is the most likely one.

2 Interpretation and Generation

The two major aspects of dialogue management are the interpretation of incoming (user) utterances, and the timely and appropriate generation of utterances by the dialogue system. To cover these aspects we have constructed a Bayes Net as shown in Figure 1.

¹This research was (partially) funded under the Wallenberg laboratory for research on Information Technology and Autonomous Systems (WITAS) Project, Linköping University, by the Wallenberg Foundation, Sweden.

In the implementation of this network we used CIspace’s Java-based Bayes Net toolkit.²

The conditional probability table for the nodes representing the dialogue move at time t and salience list at time t are obviously the core of the network. These tables are too large to present in this paper. We constructed them by hand, using heuristics gained from experience in programming rule-based dialogue systems. In future, the tables could be learned from data, or could instantiate continuous-valued functions of rule-based systems.

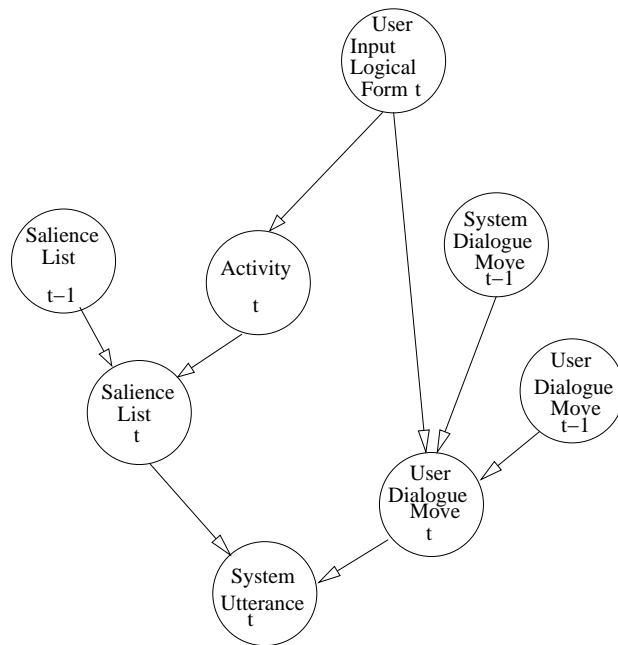


Figure 1: A Prototype Bayes Net for dialogue management

3 Anaphora resolution

Several different factors enter into the resolution of anaphora in a dialogue. How recently a potential referent was referred to is one important factor, another is the embedding activity within which the anaphoric reference was made (e.g. the type of verb phrase in which the referent appears), a third is the intra-sentential location of the relevant noun phrases in the preceding dialogue, a fourth is the relative prominence of a potential referent in the dialogue situation, and so on. The basic idea is that condi-

tional probability distributions are generated dynamically in the Bayesian network. When we look at a distribution corresponding to a node we are interested in, then the most salient object in the context will be the one whose value has the highest probability.

We point out that an obvious deterministic way to rank different combinations of factors (in an optimality-theoretic way, for example), and choose the most salient object based on this ranking, does not seem to work – because any potential ranking of principles (e.g. “Recency overrides subject placement”) would have to be context-dependent, and this defeats the idea of a ranking. See the examples in Figures 5 and 6.

3.1 Examples

Here we work with two basic activities of the WITAS robot helicopter (see (Lemon et al., 2002)) which our dialogue system interfaces to – moving and searching. The helicopter can move to various landmarks (e.g. the tower, the church) and waypoints, and can see various objects (e.g. landmarks, vehicles). Our results use the network in Figure 1. In reading the tables (the figures appear after the references), use the following key:

U =user, S =system, r =red car, g =green car, w =waypoint, s = search, m =move. All examples start with an even distribution of 0.2 for all variables (all objects are equally salient) at the start of each dialogue.

3.1.1 Activity and recency

We will start with what may be the simplest type of example of anaphoric reference, in Figure 2. Here it is intuitively clear that “it” is intended to pick out the green car. The contribution of “see it” is modelled as an observed even distribution over all possible referents which can be seen (i.e. rs and gs each get a 0.5 weight). The conditional probability table for Saliency List at time t is then used to compute the new probability distribution over the object-activity pairs (rm, gm, wm, rs, gs). Here we see that the green car is the most salient after the user’s second utterance (gm), and that this salience increases after the utterance “it”, because g was both the most recent NP, and is also a possible object in the context of the “seeing” activity.

²www.cs.ubc.ca/labs/lci/CIspace/

In the example in Figure 3, the anaphoric pronoun “it” should pick out the red car and not the waypoint, even though the waypoint was referred to more recently. Intuitively, this is because the embedding activity of looking for the red car is tied to the pronoun, and this fact overrides the most recent referent. Here, the waypoint is not a possible object in the “seeing” activity, whereas the red car has been introduced as part of that activity. Thus the pronoun “it” in the user’s final utterance has the effect of raising the probabilities of all the objects which can be seen, and this in fact overrides the recency effect of the utterance of “waypoint”.

An extended example (not shown) shows how activity information can outweigh recency in an interleaved fashion and then that a newly introduced referent can become the most salient. Having considered the ways in which activity and recency interact in determining salience for anaphoric resolution, we then investigated adding another determining factor in the model – the syntactic placement of the referring expression.

3.1.2 Placement, activity, and recency

Figure 4 shows how subject placement influences availability for anaphoric reference. Here, the subject (“red car”) of the user’s second utterance is intuitively the one picked out by the later anaphoric expression, and not the green car, even though “the green car” is the most recent NP. See Figure 4 for our results, using an extension of the network in Figure 1, where the “Activity t ” node was enriched to include syntactic information about the input – specifically, what referring expressions appear in subject and object places. Note here that the red car becomes the most salient object after the user’s second utterance. We model the referential import of this sentence as an input triple “ rng ” to the Activity t node – denoting: red car (subject), no activity, green car (object). The updated table for this node ensures that objects in subject place are given more weight than those in object place.

In Figure 5, the subject (“red car”) of the user’s second utterance is intuitively the one picked out by the later anaphoric expression, and not the green car, even though “the green car” is involved in the “seeing” activity.

In Figure 6 the red car is most salient after the

second utterance, but the waypoint becomes more salient, even though the red car was in subject position, because the waypoint is involved in the activity of moving, as is the pronoun “it”, and so is a better candidate for anaphoric resolution. Combined with Figure 5 this shows that no static ranking of anaphoric binding principles will cover all situations, and that a probabilistic approach is useful – even as a theoretical model.

Obviously this model could be made more complex with representations for direct and indirect objects, and so forth, but we leave this for future work.

4 Conclusion

We presented a Bayes Net which we have implemented to deal with dialogue move interpretation and reference resolution, and gave examples of its use for weighing a variety of factors (recency, activity, placement) in anaphoric resolution in particular. We saw that many of the insights of deterministic approaches (e.g. the WITAS Project, see (Lemon et al., 2002)) can be preserved, often in less brittle forms, in a probabilistic setting. We also have unpublished results for dialogue move classification.

References

- Simon Keizer. 2001. A probabilistic approach to dialogue act clarification. In *Proceedings of Bi-Dialog 2001*.
- Oliver Lemon, Anne Bracy, Alexander Gruenstein, and Stanley Peters. 2001. Information states in a multimodal dialogue system for human-robot conversation. In Peter Kühnlein, Hans Reiser, and Henk Zeevat, editors, *5th Workshop on Formal Semantics and Pragmatics of Dialogue (Bi-Dialog 2001)*, pages 57 – 67.
- Oliver Lemon, Alexander Gruenstein, Alexis Battle, and Stanley Peters. 2002. Multi-tasking and collaborative activities in dialogue systems. In *Proceedings of 3rd SIGdial Workshop on Discourse and Dialogue*, Philadelphia. (to appear).
- Prashant Parikh. 1990. Situations, games, and ambiguity. In R. Cooper, K. Mukai, and J. Perry, editors, *Situation Theory and its Applications I*. CSLI Publications.
- Prashant Parikh. 2001. *The Use of Language*. CSLI Publications, Stanford, CA.
- Stephen G. Pulman. 1996. Conversational games, belief revision and bayesian networks. In *7th Computational Linguistics in the Netherlands (CLIN) meeting*.

Utterance	P(rm)	P(gm)	P(wm)	P(rs)	P(gs)	Salient
U: Search for the red car S: Okay	.22	.06	.06	.6	.06	red car
U: Go to the green car S: Okay	.066	.53	.018	.18	.206	green car
U: Can you see it?	.0196	.7002	.0054	.054	.2206	green car

Figure 2: Probability distributions in example: Recency

Utterance	P(rm)	P(gm)	P(wm)	P(rs)	P(gs)	Salient
U: Do you see the red car? S: No	.18	.06	.06	.64	.06	red car
U: Go to the waypoint S: Okay	.054	.018	.718	.192	.008	waypoint
U: Do you see it?	.1108	.1036	.2154	.313	.2572	red car

Figure 3: Probability distributions in example: Activity overrides recency

Utterance	P(rm)	P(gm)	P(wm)	P(rs)	P(gs)	Salient
U: Go to the green car S: Okay	.06	.524	.06	.06	.296	green car
U: The red car is beside the green car S: Okay	.333	.1899	.018	.333	.1261	red car
U: Can you see it?	.1955	.1622	.0054	.3543	.2825	red car

Figure 4: Probability distributions in example: Subject placement overrides recency

Utterance	P(rm)	P(gm)	P(wm)	P(rs)	P(gs)	Salient
U: Search for the green car S: Okay	.06	.296	.06	.06	.524	green car
U: The red car is beside the green car S: Okay	.263	.2006	.018	.263	.2554	red car
U: Can you see it?	.1796	.1622	.0054	.3282	.3206	red car

Figure 5: Probability distributions in example: Subject placement overrides activity

Utterance	P(rm)	P(gm)	P(wm)	P(rs)	P(gs)	Salient
U: Go to the waypoint	.06	.06	.76	.06	.06	waypoint
R: The red car is at the way-point	.333	.018	.298	.333	.018	red car
U: Did you get to it?	.2888	.1582	.3288	.1444	.0913	waypoint

Figure 6: Probability distributions for example: Activity overrides subject placement