A Computational Architecture for Conversation

Eric Horvitz[1] and Tim Paek[2 *]

[1] Decision Theory and Adaptive Systems, Microsoft Research, WA, U.S.A.
[2] Department of Psychology, Stanford University, CA, U.S.A.

Abstract. We describe representation, inference strategies, and control procedures employed in an automated conversation system named the *Bayesian Receptionist*. The prototype is focused on the domain of dialog about goals typically handled by receptionists at the front desks of buildings on the Microsoft corporate campus. The system employs a set of Bayesian user models to interpret the goals of speakers given evidence gleaned from a natural language parse of their utterances. Beyond linguistic features, the domain models take into consideration contextual evidence, including visual findings. We discuss key principles of conversational actions under uncertainty and the overall architecture of the system, highlighting the use of a hierarchy of Bayesian models at different levels of detail, the use of value of information to control question asking, and application of expected utility to control progression and backtracking in conversation.

1 Introduction

Conversations are initiated to express needs or to acquire, share, or critique information. Ongoing inference about intentions is critical in conversation (Allen and Perrault, 1980). We explore the role of inference and decision making under uncertainty in conversational dialog. People engaged in conversation appear to make inferences under uncertainty about the relevance of multiple sources of information, including nonlinguistic contextual cues such as visual findings. They also make ongoing decisions about the formulation of discriminating questions and about the most appropriate level of detail at which to exchange information. We describe in this paper methods for modeling and automating conversation that leverage inference, decision making, and information gathering under uncertainty. We present a computational architecture that focuses on issues of representation, uncertain inference, and dialog control procedures for navigating among different levels of detail in conversation. We highlight basic principles in the context of an implementation of an automated conversation prototype developed at Microsoft Research named the *Bayesian Receptionist*.

The Bayesian Receptionist prototype represents a melding of several key components. Bayesian inference is employed at distinct levels of an abstraction hierarchy to infer a user's goals. Value of information procedures are employed within levels to gather additional information, and decision-theoretic strategies using threshold analyses guide the progression and backtracking in conversation based on expected utility. The system considers words and phrases as well as higher-level linguistic distinctions obtained via a natural language parse of a user's utterances. The un-

* We are grateful for assistance provided by Mike Barnett, Herb Clark, and Andy Jacobs.

certain reasoning machinery drives a social user interface built with the MS Agent package. The interface provides speech recognition, speech synthesis, and animations that provide an anthropomorphic presence.

2 The Receptionist Domain

The Bayesian Receptionist project centers on conversation about goals typically handled by receptionists at the front desks of buildings on the Microsoft corporate campus. Receptionists serve a vital role in facilitating daily activities at each building at Microsoft. When people interact with the receptionist, they engage in a *joint activity* (Clark, 1996; Jennings and Mamdani, 1992; Levinson, 1992). A joint activity is a task-oriented, social event with constraints on participants, setting, and most of all, reasonable or allowable contributions. Participants in a joint activity assume that they share some common set of beliefs about the activity, including assumed roles and responsibilities with other participants and the degree to which participants are attending to and understanding the content of utterances (Paek and Horvitz, 1999)

We conducted an observational study of the receptionist domain by videotaping nine hours of participants interacting with three receptionists. Through reviewing videotapes and engaging receptionists in discussions, we identified a key set of user goals, as well as key pieces of linguistic and visual information relevant to the problem of diagnosing a user's needs. Discussion with the receptionists revealed 32 mutually exclusive and exhaustive goals that make up the joint activity. The videotapes and interviews also revealed critical variables and states considered by receptionists. For example, we found that receptionists reason about the *type* of person (*e.g.*, a person visiting the campus to interview for a position). We identified 12 types of people and assessed links between person type and variables capturing appearance and patterns of behavior, as well as the different prior probability distribution over goals for people in different classes. Visual information employed by receptionists in making inferences about goals include:

- **Appearance**: the attire and type of identification badge that may be visible.
- **Behavior**: whether the user looks hurried or glances outside during the interaction.
- **Spatial configuration**: the mode of arrival and the trajectory of locomotion in the proximity of the receptionist.
- **Props**: whether the user carries some type of equipment, or was in a group of people.

In addition to visual features, we identified a set of initial utterances associated with different goals. Analysis of real-world interactions with the receptionist showed considerable linguistic variability. At times, people employed conventional phrasing such as "I'm here to see Rick Rashid." However, people frequently interacted with more telegraphic, abbreviated utterances such as "Bathroom?" to communicate the goal, "I have to use the bathroom; where is it?" and "18!" (while running toward the building exit) to relay the urgent need to order a shuttle heading to Building 18. Others have noted such use of succinct utterances in combination with contextual cues (Clark, 1996). We also found that people often used creative ad-hoc phrases, as in, "Beam me to 25 please" to indicate a need for a campus shuttle to Building 25. We even observed "contextual constructions" (Clark, 1983; Nunberg, 1979) as in "Do we do Sea-Tac airport?"

Considerable variability in dialog is possible since people take it for granted that participants in a conversation will be able to infer what they mean from their shared beliefs. We found that recep-

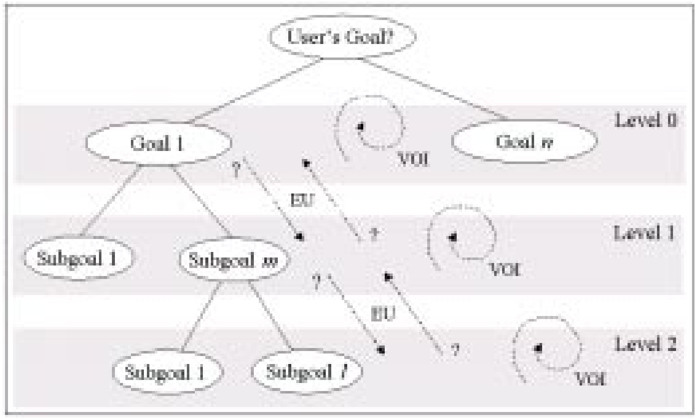

Figure 1. Schematized view of an overall model for guiding diagnosis of goals via Bayesian inference, computation of value of information (VOI), and expected utility analysis (EU) for navigation between levels of detail for progression and backtracking in conversation.

tionists are very apt at discerning what people need, relying on shared knowledge about the joint activity and the likely goals of the participants to guide the dialog. Without such knowledge it would be difficult, if not impossible, for people to communicate their goals with receptionists as efficiently as they do. Also, it became apparent that inferences about the goals of people seeking assistance depend upon integrating rich patterns of evidence beyond the utterance. Visual information, including the appearance, trajectory, and behavior of the participants, is often taken into account.

3 Hierarchical Decomposition of Conversation about Goals

Our observations of interactions, coupled with attempts to model and automate dialog as decision making under uncertainty, led us to focus on the use of a task abstraction hierarchy as an organizing representation. The task abstraction hierarchy decomposes the problem of understanding a user's goals into diagnoses at successive levels of detail. Figure 1 displays the task abstraction hierarchy used in the Bayesian Receptionist. Level 0 represents the task of discriminating the high-level goal of the user given initial observations and an initial utterance. Level 0 goals include the need to arrange a shuttle to travel to another part of campus, to enter the building, to pick something up, to drop something off, to get directions, to acquire an answer to an informational query, to acquire help with a special project, to pick up a public commuting pass, to send a fax, and to hang posters or remove posters. In order to render the set of goals exhaustive, we introduced a goal labeled *other* to refer to goals that are not explicitly considered in the high-level goal variable. Level 1 represents the refinement of the high-level goals into more specific needs. For example, the Level 1 refinements of the goal of needing a shuttle to travel to another part of campus include the need to travel as a group or as a single person to a main campus location, to the North Campus (MS Interactive Media Division), and to a non-Microsoft location. Level 2 considers the additional specification of types of shuttles for special cases, including the need for a shuttle equipped for transporting handicapped people and for express shuttles to transport senior executives. Levels more detailed than the highest level include an additional state representing the proposition that the current level is inappropriate. Inference about the belief assigned to this state is used to control backtracking in conversation.

We found several advantages to introducing a goal refinement hierarchy rather than attempting to model the problem of goal understanding to a larger problem at a single level of detail. Decomposition of a user's goals into several levels of detail allows for guiding conversation on a natural path of convergence toward shared understanding at progressively greater detail. Multiple levels also allow for the establishment of *common ground* (Clark, 1996) about uncertainties at each level and for conversation *about* comprehension of misunderstandings before progressing to the next le-

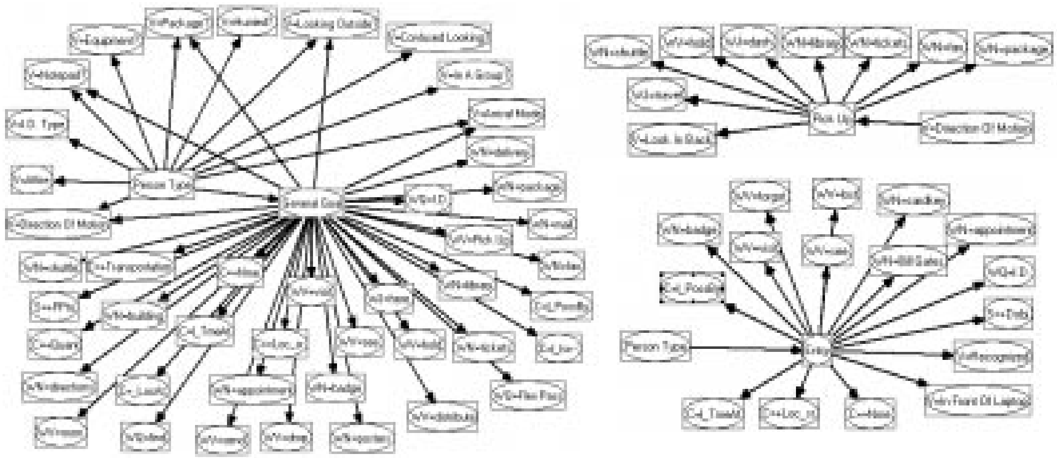

Figure 2. Bayesian networks for inference about primary high-level goals (left) and sample models for the intentions of gaining access to the building and picking up materials stored at the front desk (right).

vel of detail. That is, users can be directed, as part of a natural dialog about their goals, to implicitly or explicitly confirm or disconfirm misunderstanding at one level, reducing uncertainty to some tolerable level before progressing to the next level. Our observations suggested that such discussion about different levels of detail is common and often expected as part of natural human—human conversation. From an engineering and computational perspective, decomposing the goal understanding task into a set of nearly decomposable understanding subproblems leads to more tractable, level-specific modeling and inference. However, the introduction of multiple subproblems also introduces the need for elegant control of inference, evidence gathering, and decision making within and between the levels.

In real-time use, we employ the hierarchical decomposition to disambiguate goals by asking the most informative questions or gathering the most valuable non-linguistic evidence to identify goals at progressively more detailed levels. Given confirmation or strong belief in a speaker's goal at the current level, the system passes all of its evidence to the next more detailed level of analysis and attempts to refine its beliefs about goals at that level, until reaching a conclusion about the speaker's ultimate goals. In Sections 4 through 8, we describe details on inference, information gathering, and control policies for navigating among multiple levels of analysis.

4 Bayesian Models for Inferring a User's Communication Goals

We worked to build models for inference and decision making under uncertainty about a user's goals at each level of the task hierarchy with Bayesian networks. Bayesian networks have been useful in representing a variety of challenges centering on user modeling under uncertainty (Albrecht et al., 1997; Conati et al., 1997; Horvitz and Barry, 1995; Horvitz et al., 1998; Jameson, 1996). Bayesian-network models allow a system to fuse together the relevance of multiple classes of information allowing for inference about the goals of a user from visual information and utterances. Also, having access to probabilities at any step in a dialog allow us to guide question asking based on the computation of value of information.

We constructed and assessed by hand Bayesian networks considering linguistic and non-linguistic observations. The Bayesian networks were authored in the Microsoft Research MSBN modeling and inference system. Models were built for different levels of the task hierarchy. The non-linguistic observations in the models bring to bear contextual information, including important visual features about the appearance, location, and trajectory of locomotion of people interacting with the system. Figure 2 displays the structure of Bayesian models for the Level 0 discrimination

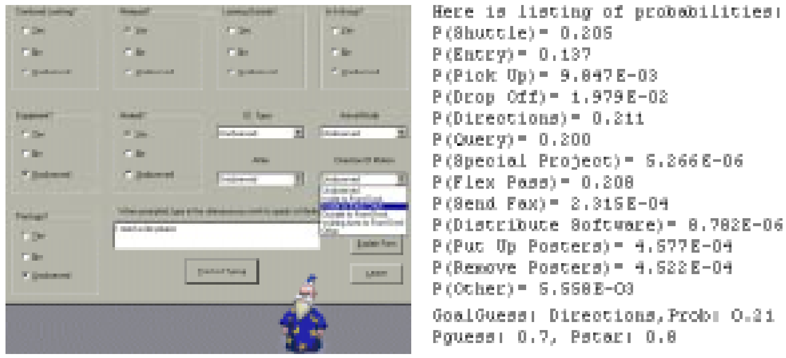

Figure 3. Bayesian Receptionist screen displaying a query initiating a conversation and the probability distribution over Level 0 goals inferred with a Bayesian model taking into consideration linguistic and visual cues.

problem and for two of the Level 1 subproblems centering on the goals of picking something up from the receptionist and gaining entry to the building.

In use, Bayesian networks at progressively more detailed levels of analysis are passed all of the linguistic and non-linguistic information that has been observed by the system. Figure 3 shows an initial screen of the Bayesian Receptionist considering the utterance, “I need a ride please.” In addition to the linguistic information, the system considers the visual findings that the user appeared hurried, carried a notebook, and approached the automated receptionist with a trajectory characterized as “moving from inside the building to the back of the receptionist desk.” The probability distribution over goals at Level 0 inferred from the visual observations and linguistic clues is displayed to the right of the screen.

5 Accessing Linguistic Evidence with Natural Language Processing

Introducing linguistic distinctions into the Bayesian models and making linguistic observations at run time posed a challenge. The taped interactions with the receptionist showed that utterances varied in syntax, length, and typicality. We first worked to identify evocative sets of words and phrases spotted in utterances, referred to as *metanyms* (Heckerman and Horvitz, 1998). To extend the ability to handle the variation in feasible linguistic inputs to the Bayesian models, we employed automated natural language processing (NLP) to also introduce higher-level linguistic abstractions as observational variables in our Bayesian models. We submitted sample utterances for different goals to a system named NLPwin (Heidorn, 1999), developed by the MS Research Natural Language Processing group to identify *syntactic*, *logical*, and *semantic* cues that can be useful for distinguishing between goals. After submitting myriad utterances of different length, typicality, and syntactic structure to NLPwin, we selected a set of frequently occurring, discriminatory cues as evidential variables in the Bayesian networks.

The textbox at the center of Figure 3 shows the user submitting the utterance, “I need a ride please.” The sentence is processed in five stages of analysis. In the first stage, the system segments the input utterance into individual tokens, analyzing words by their morphological structure and looking them up in online dictionaries, including sources specifically tailored for multi-word entries. In the second stage, known as the syntactic sketch, the system parses the utterance into its syntactic components based on rules of grammar. In the third stage, known as the syntactic portrait, the system resolves attachment ambiguities by using semantic relations culled from a compiled analysis derived from online dictionary definitions. The result of the first three stages is the parse tree shown in the top panel of Figure 4. A fourth stage resolves anaphoric references and constructs the logical form of the sentence, representing predicate-argument relations in a semantic graph, by assigning sentence elements to functional roles, such as deep subjects in a semantic graph, by assigning sentence elements to functional roles, such as deep subjects (Dsub)

```
I need a ride please.
DECL1     NP1       PRON1*    "I"
          VERB1*    "need"
          NP2       DETP1     ADJ1*     "a"
                    NOUN1*    "ride"
          AVP1      ADV1*     "please"
          CHAR1     "."
  ------------------------------
need1 (+Pres +T1)
|_Dsub-----I1 (+Pers1 +Sing +Anim +Humn)
|_Dobj-----ride1 (+Indef +Pers3 +Sing +Count)
|_Mods-----please1 (+FO)
  ------------------------------
2;need;VERB;0;0;-1;;; +Pres +T1 +Economics +Sociology +Economics +So
1;I;PRON;0;0;-1;;; +Pers1 +Sing +Anim +Humn
4;ride;NOUN;0;0;-1;;; +Indef +Pers3 +Sing +Count +EQ +SP +EQ +TN +ON
5;please;ADV;2;Mods;-1;;; +FO
  ------------------------------
    C=_LocAt@   C=+Loc_sr@  C=+Kme@ C=+Quant@   C=+Transportation@
```

Figure 4. Output of the natural language analysis of the query, "I need a ride please," yielding syntactic, logical, and semantic clues.

and deep indirect objects (Dind). The second panel of the output displays the logical form, concluding that the deep object (Dobj) is the indefinite, count form of the argument "ride." The final stage of NLPwin attempts to determine the most appropriate sense for words in the utterance from a list of senses, yielding semantic cues for the Bayesian models. These are listed in the third panel of the NLP analysis. As an example, an inferred sense of the noun "ride" is "+TN", short for Transportation. Linguistic distinctions noted in the parse, including the identification of semantic concepts of transportation and location are considered in the Level 1 Bayesian network.

6 Value of Information to Guide Observation and Dialog

A critical component of decision-theoretic diagnosis (Horvitz et al., 1988) is the identification of the most valuable additional observations to collect to enhance the value of actions ultimately taken in the world. Formalisms for identifying the most valuable information to collect under uncertainty hold promise for guiding automated dialog (Jameson, 1995). We have employed *value of information* (VOI) analysis to identify the best questions to ask and visual observations to make in light of the inferred probabilities of different goals at distinct levels of the goal decomposition hierarchy. VOI yields the expected utility of evaluating different pieces of previously unobserved evidence, considering the informational value and the cost of making the observation under uncertainty. To compute VOI, we consider, for each observation, the expected utility of the best decision associated with each value that the observation may take on. The analysis sums the expected utility for each value of future observation, weighted by the probabilities of seeing the different values, should that observation be made.

An exact computation of VOI requires the consideration of all possible sequences of observations. However, greedy VOI, focusing on computing the next best single piece of evidence to observe, has been found to be a good approximation (Gorry and Barnett, 1968). Within the framework of greedy VOI, a variety of approximations have been employed. Beyond an explicit decision-theoretic computation, investigators have explored an information-theoretic version of VOI based on the minimization of entropy (Gorry and Barnett, 1968; Heckerman et al., 1992), and the use of the statistical properties of large samples to develop VOI approximations.

In the Bayesian Receptionist, we harness information-theoretic VOI analysis to control information gathering to resolve the current uncertainty within levels of analysis. This approach is related to prior work on the use of VOI at multiple levels of abstraction in decision-theoretic diagnostic systems (Horvitz et al., 1989). After each new observation is evaluated, the system updates the probabilities of distinct goals within a level of the task hierarchy and recomputes the VOI. VOI continues within a level until either the expected cost of evaluating observations exceeds the expected value of the observations, there are no additional observations to make, or a higher-level decision-theoretic analysis progresses the system to the next level of detail. We now discuss the control of progression between adjacent levels of the refinement hierarchy.

7 Progression to Additional Levels of Detail

Given strong belief in a goal at some level of detail, it is typically appropriate to move on to perform inference and question asking about goals at a more detailed level of analysis. Our intuitions about the progression to additional levels of detail were developed by considering the ways people in a conversation in the context of a joint activity speak in a natural manner *about* goals and subgoals (Cohen and Levesque, 1994). At any point in a conversation, a participant may assume a goal or ask for direct confirmation about a goal or subgoal. A correct explicit guess can raise the efficiency of communication. However, a poor guess can be costly. Poor guesses, especially when coming early in a conversation, may appear unnatural and can relay an impression that the listener is rushing or is simply not considering obvious clues. Instead of communicating a guess, an agent may simply decide to assume a goal based on the probabilities of alternate goals at a level and progress to a more detailed level of discourse. Although correct action can lead to efficient conversation, erroneous decisions to progress are usually even more costly than explicit guessing as users are not given the chance to converse about the decision. We can consider such costs and benefits of alternate progression actions with a decision-theoretic analysis.

We developed and implemented methods for performing cost–benefit analysis to guide decisions about progression to the next level of detail, or to return to a previously examined level, based on inferred probabilities of goals within a level and the associated costs and benefits of progressing. As highlighted in Figure 1, at every step of the analysis, we consider the expected value of directly progressing or asking a question to confirm a goal, versus remaining at the current level of analysis and continuing to gather information based on VOI. We assess the utilities of different conversational outcomes and employ an approximate decision analysis to compute threshold probabilities for progression, seeking confirmation, or backtracking to a previously visited level. Details of the derivation of probability thresholds from a cost—benefit analysis are described in (Horvitz, 1999). Beyond derivation via consideration of the utility of outcomes, such threshold probabilities can be assessed directly.

The Bayesian Receptionist makes use of three thresholds, *p* progress, p* guess,* and *p* backtrack*. If the probability of the goal with the highest likelihood does not exceed *p* guess* or *p* progress,* the system continues to perform VOI to gather additional information about the goals. If VOI becomes non-positive or there are no other observations to make, the system issues a request for additional information. Should the maximal likelihood goal at the current level exceed *p* progress*, the system will assume the goal and move to the next level of detail. Should the maximal likelihood exceed *p* guess*, but not exceed *p* progress*, the system will engage the user to confirm or rule out the goal. Following a crisp disconfirmation of a goal, a probability of zero is assigned to that goal and the probability distribution for the remaining feasible goals is renormalized. The *p** threshold for *backtracking* is used in making a decision to return to a higher level of detail. The Bayesian models for goal disambiguation at levels more detailed than Level 0 include a *return* hypothesis. Should the probability of all subgoals decrease and the likelihood of the return hypothesis become greater than *p* backtrack,* the system returns to the previous level. In such situations, the system apologizes for the misunderstanding and either continues VOI at the previous level or asks for more information if that VOI is non-positive.

Beyond the use of static utilities in the computation of the decision thresholds and VOI, we employed context-sensitive utilities. Utilities used in the computation of the threshold probabilities and VOI can be defined as functions of such observations as the user being hurried, or of observations about the conversational dialog itself such as the number of questions that have already been asked of users. The Bayesian Receptionist makes use of a special class of probability-sensitive costs for controlling VOI when deliberating about asking users questions at levels of detail that are marked as *atypical* levels of analysis. An example of an atypical level of analysis is Level 2, associated with the high-level goal of desiring a shuttle to travel somewhere on the Microsoft campus. Level 2 seeks to discover whether a person will need a special kind of shuttle, including the need for a vehicle that can transport handicapped people or a special shuttle for senior executives. Levels marked as atypical include a special goal state that represents the proposition that the level is irrelevant. Engaging users about subgoals at this level would appear unnatural in most cases. We dynamically update the expected cost used in the VOI analysis by considering the probability that an atypical level is relevant. The probability of relevance given evidence, $p(R|E)$, is the complement of the probability assigned to the state *irrelevent*. As the probability of relevance increases, the naturalness of being asked increases, making such a question less surprising, and, thus, less costly to ask. We compute the expected cost of asking as the sum of the costs for the outcomes of the level being relevant and irrelevant, weighted by the inferred probabilities as follows:

$$\text{Expected Cost} = p(R\,|E)\ \text{Cost}(\textit{Asking}, \textit{Relevant}) + 1\text{-}p(R\,|E)\ \text{Cost}(\textit{Asking}, \textit{not Relevant})$$

where Cost(Asking, Relevant) is the cost of asking a question about an atypical situation when the level is relevant and Cost(Asking, not Relevant) is the cost of asking when the level is irrelevant. We use this expected cost in VOI for explicit question asking (versus those centering on the gathering of visual clues) in atypical levels of detail.

8 A Sample Conversation

Before concluding, let us complete the conversation started with the Bayesian Receptionist in Figure 3. As the dialog progresses, the system asks for additional information and notices the evocative word “shuttle” in the user’s utterance. Given this observation, the probability that the user’s goal is to travel somewhere by shuttle rises to 0.69. The belief in the leading hypothesis is not high enough to reach *p* guess* or *p* progress* so VOI is called. The system continues to gather evidence until the probability of the leading goal surpasses *p* guess*. At this point, the system seeks direct confirmation, asking, “So you’d like a shuttle?” After confirmation, the system progresses to Level 1, focused on disambiguating the destination of the user and whether the user is traveling alone or in a group. The linguistic observation, “north” is noticed and added to the analysis and the probability distribution is updated. The greatest likelihood is now assigned to the goal of traveling to the North Campus alone, followed by travel to the North Campus in a group. However, the probabilities do not exceed *p* guess*, so the system calls VOI. As portrayed in the upper portion of Figure 5, the VOI analysis leads to a decision to seek assistance with evaluating visual information about the presence of a group of people. The agent “observes” a group of people and integrates this evidence to update the probability distribution displayed in the center of the upper panel. The probability of this goal exceeds *p* guess*, but does not exceed *p* progression*, so the

Figure 5. Upper sequence: Use of VOI to acquire visual information confirming the presence of a group, followed by update of probabilities for Level 1 goals and a decision to confirm a goal explicitly. Lower sequence: Progression to next level, use of VOI to detect if user is a recognized executive, update of probabilities of goals, and decision to forego interaction with the user about an atypical condition.

system seeks confirmation of the goal at this level of detail. Progression now occurs to Level 2 and inference is called followed by VOI. This level is labeled as an *atypical* level of refinement. Given the atypia of the level, the system computes the probability-dependent expected costs of asking the user questions and uses these costs to modulate the cost-benefit analysis in the VOI computation. Explicit questions about whether the user is an executive or is handicapped are suppressed, and the system decides to simply examine the situation with vision. As captured by the lower panel in Figure 5, visual information is acquired. The user is not recognized as an executive and inference is called. The maximal likelihood subgoal is *normal*, referring to the irrelevance of the atypical level of detail. The inferred probability of normal surpasses the p^* threshold and the system confirms its understanding of the user's goal at Level 1, completing the conversation.

9 Summary and Ongoing Work

We introduced an architecture for modeling dialog in joint-activity situations. Central to the architecture is the use of a goal decomposition hierarchy employing Bayesian models at increasingly detailed levels of analysis and the use of cost—benefit control procedures for navigating between levels of the hierarchy. The run-time prototype relies on the integration of several inferential components including natural language processing to provide a stream of linguistic evidence, Bayesian inference to update beliefs about the set of goals at distinct levels of analysis, value of information for asking the best questions or making the most informative observations, and decision-theoretic control policies for determining when to progress to the next level of analysis or to halt with an action. We described the operation of a prototype when populated with models built through consideration of the real-world tasks faced by receptionists at Microsoft. In our continuing work, we are pursuing the use of richer control strategies, learning of models from collected data, and extending the analysis to consider the status of a user's attention and of the mutual understanding of the semantics of concepts being communicated in conversations.

References

Albrecht, D.W., Zukerman, I., Nicholson, A.E., Bud, A. (1997). Towards a Bayesian model for keyhole plan recognition in large domains. In Jameson, A, Paris, C., and Tasso, C., eds., *Proceedings of the Sixth International Conference on User Modeling*. New York: Springer-Verlag. 365-376.

Allen, J.F., and Perrault, C. (1980). Analyzing intention in utterances. *Artificial Intelligence* 15:143-178.

Clark, H.H. (1983). Making sense of nonce sense. In G.B. Flores d'Arcais and R. Jarvella, eds., *The Process of Language Understanding*. New York: Wiley. 297-331.

Clark, H.H. (1996). *Using Language*. Cambridge University Press.

Cohen, P.R., and Levesque, H.J. (1994). Preliminaries to a collaborative model of dialogue. *Speech Communication* 15:265-274.

Conati, C., Gertner, A.S., VanLehn, K., Druzdzel, M.J. (1997). Online student modeling for coached problem solving using Bayesian networks. In Jameson, A, Paris, C., and Tasso, C., eds., *Proceedings of the Sixth International Conference on User Modeling*. New York: Springer-Verlag. 231-242.

Gorry, G.A. and Barnett, G.O. (1968). Experience with a model of sequential diagnosis. *Computers and Biomedical Research* 1:490-507.

Heckerman, D.E., Horvitz, E., and Nathwani, B.N. (1992). Toward normative expert systems: Part I. The Pathfinder project. *Methods of Information in Medicine* 31:90-105.

Heckerman, D., and Horvitz, E. (1998). Inferring informational goals from free-text queries: A Bayesian approach, *Fourteenth Conference on Uncertainty in Artificial Intelligence*, San Francisco: Morgan Kaufmann Publishers, 230-237. http://research.microsoft.com/~horvitz/aw.htm

Heidorn, G. (1999) Intelligent writing assistance. In Dale, R., Moisl, H., and Somers, H. eds., *A Handbook of Natural Language Processing Techniques*. Marcel Dekker.

Horvitz, E. Heckerman, D.E., Ng, K. and Nathwani, B.N. (1989). Heuristic abstraction in the decision-theoretic Pathfinder system, *Proceedings of the Thirteenth Symposium on Computer Applications in Medical Care*. IEEE Computer Society Press. 178-182.

Horvitz, E., Breese, J., and Henrion, M. (1988). Decision theory in expert systems and artificial intelligence. *International Journal of Approximate Reasoning, Special Issue on Uncertainty in Artificial Intelligence* 2:247-302. http://research.microsoft.com/~horvitz/dt.htm

Horvitz, E., Barry, M. (1995). Display of information for time-critical decision making. In Besnard, P., and Hanks, S., eds., *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. San Francisco: Morgan Kaufmann. 296-305. http://research.microsoft.com/~horvitz/vista.htm

Horvitz, E., Breese, J.S., Heckerman, D., Hovel, D., Rommelse, K. (1998). The Lumiere Project: Bayesian user modeling for inferring the goals and needs of software users, *Fourteenth Conference on Uncertainty in Artificial Intelligence.* San Francisco: Morgan Kaufmann Publishers. 256-265. http://research.microsoft.com/~horvitz/lumiere.htm

Horvitz, E. (1999). Principles of mixed-initiative user interfaces, In *Proceedings of Computer-Human Interaction '99*, Association for Computing Machinery Press.

Jameson, A., Schafer, R., Simons, J., and Weis, T. (1995). Adaptive provision of evaluation-oriented information: Tasks and techniques. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. 1886-1895.

Jameson, A. (1996) Numerical uncertainty management in user and student modeling: An overview of systems and issues. *User Modeling and User-Adapted Interaction* 5:193-251.

Jennings, N.R., and Mamdani, E.H. (1992). Using joint responsibility to coordinate collaborative problem solving in dynamic environments. *Proceedings of the Tenth National Conference on Artificial Intelligence*. Menlo Park: AAAI Press. 269-275.

Levinson, S.C. (1992). Activity types and language. In P. Drew and J. Heritage, eds., *Talk at Work.* Cambridge University Press. 66-100.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3:143-184.

Paek, T., and Horvitz, E. (1999), A layered representation of uncertainty for managing dialogue. Manuscript submitted for publication.