Extending the Knowledge-Level Approach to Planning for Social Interaction

Extending the Knowledge-Level Approach to Planning for Social Interaction

Ronald P. A. Petrick
School of Informatics
University of Edinburgh
Edinburgh EH8 9AB, Scotland, UK
rpetrick@inf.ed.ac.uk

Abstract

A robot coexisting with humans must not only be able to perform physical tasks, but must also be able to interact with humans in a socially appropriate manner. We describe an application of planning to task-based social interaction using a robot that must interact with multiple human agents in a simple bartending domain. Extensions to previous work include a new domain supporting planning for group interactions, and reasoning about uncertainty due to automatic speech recognition.

Introduction

As robots become integrated into daily life, they must increasingly deal with situations in which *socially appropriate interaction* is vital. In such settings, it is not enough for a robot simply to achieve its task-based goals; instead, it must also be able to satisfy the social goals and obligations that arise through interactions with people in real-world settings. To address this challenge, we are investigating *task-based social interaction* in a bartending domain, by developing a robot bartender (Figure 1) that is capable of dealing with multiple human customers in a drink-ordering scenario.

Key to our approach is the use of a high-level planner, which is responsible for action selection and reasoning in the robot system. Specifically, we use the knowledge-level planner PKS (Petrick and Bacchus 2002; 2004), a choice that is motivated by PKS's ability to work with incomplete information and sensing actions: not only must the robot perform physical tasks (e.g., handing a customer a drink), it will often have to gather information it does not possess from its environment (e.g., asking a customer for a drink order). Moreover, since interactions will involve human customers, speech will be the main input modality and many of the planner's actions will correspond to speech acts, providing a link to natural language processing—a research field with a long tradition of using planning, but where general-purpose planning techniques are not the focus of mainstream study.

In this paper we highlight two extensions to our planning approach, which build on previous work (Petrick and Foster 2013): a new planning domain supporting group interactions, and a means of processing social state information for reasoning about uncertain speech hypotheses. This work is part of the JAMES project (Joint Action for Multimodal Embodied Social Systems; see james-project.eu).

Figure 1: The JAMES robot bartender

Knowledge-Level Planning with PKS

The target application for this work is a simple bartending scenario, using the robot platform shown in Figure 1. While planning offers a tool for action selection, it is only one component in a larger system that includes visual processing, speech recognition, natural language processing, and robot control (Giuliani et al. 2013). In particular, the planner is responsible for managing interactions with customers, tracking multiple drink orders, and gathering information through follow-up questions (Petrick and Foster 2013). To do this, the planner takes state reports from a state manager and selects actions to be executed on the robot. Plans are generated using PKS (Planning with Knowledge and Sensing) (Petrick and Bacchus 2002; 2004), a conditional planner that works with incomplete information and sensing actions.

Unlike many general-purpose planners, PKS works at the *knowledge level* and reasons about how its knowledge, rather than the world, changes due to action. For efficient reasoning, PKS restricts the knowledge it can represent while ensuring it is expressive enough to model many common types of information. PKS actions are described by *preconditions*, which ask questions about the planner's knowledge, and *effects*, which modify the planner's knowledge through STRIPS-like additions and deletions. E.g., Figure 2 shows a PKS action from the bartender domain. PKS constructs plans using forward search, and can build contingent plans by considering certain outcomes arising from its knowledge.

Planning for Group Interactions

To generate plans for the robot, PKS uses a symbolic domain model that includes a specification of the physical, sensory,

action *ask-drink(?a : agent, ?g : group)*
preconds: $K(inGroup(?a) = ?g) \wedge \neg K(ordered(?a)) \wedge \neg K(otherAttnReq) \wedge \neg K(badASR(?a))$
effects: $add(K_f, ordered(?a)), add(K_v, request(?a))$

Figure 2: Example PKS action in the bartender domain

and linguistic actions available to it. In previous work, the domain supported simple interactions with individual agents for ordering drinks from the robot (Petrick and Foster 2013). More recently, we have extended this domain to include new behaviours that real bartenders exhibit in natural interactions: only agents who are seeking to engage with the bartender are addressed; the bartender acknowledges all drink orders as they are given; if a group approaches the bar, the bartender takes all of the drink orders in sequence and then serves all of the requested drinks; if a new agent appears while the bartender is engaged, the agent is acknowledged, and then served after the current transaction is complete.

These behaviours are formalised in a PKS planning domain that includes actions such as: *greet(?a,?g)* (greet agent *?a* in group *?g*), *ask-drink(?a,?g)* (ask *?a* in group *?g* for a drink order), *serve(?a,?d,?g)* (serve drink *?d* to *?a* in group *?g*), *bye(?a,?g)* (end an interaction with *?a* in group *?g*), *wait(?a,?g)* (tell *?a* in group *?g* to wait), *ack-order(?a,?g)* (acknowledge *?a*'s order in group *?g*), and *ack-wait(?a,?g)* (thank *?a* in group *?g* for waiting). Most notably, all actions are based on the idea of agent groups, which affects how individual agents are served. For instance, consider the case where three agents are in the bar: *A1* and *A2* are part of a group *G1*, and *A3* is in a singleton group *G2*. Here, the planner might build the following plan for serving all agents:

wait(A3,G2),	[Tell *G2* to wait]
greet(A1,G1),	[Greet group *G1*]
ask-drink(A1,G1),	[Ask *A1* for drink order]
ack-order(A1,G1),	[Acknowledge *A1*'s order]
ask-drink(A2,G1),	[Ask *A2* for drink order]
ack-order(A2,G1),	[Acknowledge *A2*'s order]
serve(A1,request(A1),G1),	[Give the drink to *A1*]
serve(A2,request(A2),G1),	[Give the drink to *A2*]
bye(A2,G1),	[End *G1*'s transaction]
ack-wait(A3,G2),	[Acknowledge *G2*'s waiting]
ask-drink(A3,G2),	[Ask *A3* for drink order]
ack-order(A3,G2),	[Acknowledge *A3*'s order]
serve(A3,request(A3),G2),	[Give the drink to *A3*]
bye(A3,G2).	[End *G2*'s transaction]

The plan first directs the robot to tell group *G2* to wait before transacting with group *G1*. The robot then collects drink orders from all customers in *G1* before serving their drinks and completing the transaction. (The term *request(A)* acts as a placeholder for the actual drink ordered by customer *A*.) After that, the robot thanks group *G2* (i.e., customer *A3*) for waiting before taking and serving the final drink order.

Planning for Uncertain Speech Hypotheses

We are also extending our management of more complex scenarios that arise in dialogue-based interaction, by improving our ability to plan under uncertainty due to automatic speech recognition (ASR). Currently, any speech input hypotheses other than the top hypothesis (h_u^*, c_u^*) is discarded by the natural language processing module. As a result, potentially high-likelihood alternatives to the top hypothesis are unavailable to the planner, raising the possibility of less effective (or incorrect) action choices by the planner during plan construction. Instead, we are exploring the idea of passing an n-best list of processed hypotheses to the state manager for inclusion in the state representation, as a set of alternative interpretations for an agent's utterance. In practical terms, the n-best can be determined by the list of top entries that account for a significant probability mass in terms of the hypotheses' associated confidence measures, i.e., $\{\langle h_1, c_1\rangle, \langle h_2, c_2\rangle, \ldots \langle h_n, c_n\rangle\}$, such that $\sum_{i=1}^{n} c_i > \theta$, where θ is some threshold. Using this list, the state manager can then derive a set of interpretations, $\{\phi_1, \phi_2, \ldots, \phi_n\}$, where each ϕ_i is a conjunction of state fluents. (In our domain, each ϕ_i is usually a single fluent.)

At the planning level, such disjunctive state information can be represented in PKS's K_x database as an "exclusive or" formula of the form $(\phi_1|\phi_2|\ldots|\phi_n)$. Once such information is available in planner's knowledge state, it can be directly used during plan construction. In practice, such knowledge often has the effect of introducing additional sensing actions into a plan, to disambiguate between K_x alternatives. To aid this process, we are adding new domain actions which correspond to information-gathering questions that the robot can ask to help clarify uncertain beliefs (without asking an agent to simply repeat an utterance, which is often interpreted by humans as a poor dialogue move).

Conclusions

While these additions extend our bartending scenario, we also want to support more complex interactions, including agents that can ask questions about drinks or order multiple drinks. Each of these extensions will require a more active role for the planner. However, based on early results (Giuliani et al. 2013), we believe that general-purpose planning continues to offer a promising tool for action selection in task-based interaction, as an alternative to more specialised approaches used in many interactive dialogue systems.

Acknowledgements: Thanks to members of the JAMES consortium who helped implement the bartender system. This research received funding from the European Commission's 7th Framework Programme under grant no. 270435.

References

Giuliani, M.; Petrick, R. P. A.; Foster, M. E.; Gaschler, A.; Isard, A.; Pateraki, M.; and Sigalas, M. 2013. Comparing task-based and socially intelligent behaviour in a robot bartender. In *Proceedings of ICMI 2013*, 263–270.

Petrick, R. P. A., and Bacchus, F. 2002. A knowledge-based approach to planning with incomplete information and sensing. In *Proceedings of AIPS 2002*, 212–221.

Petrick, R. P. A., and Bacchus, F. 2004. Extending the knowledge-based approach to planning with incomplete information and sensing. In *Proceedings of ICAPS 2004*, 2–11.

Petrick, R. P. A., and Foster, M. E. 2013. Planning for social interaction in a robot bartender domain. In *Proceedings of ICAPS 2013, Special Track on Novel Applications*, 389–397.