# A Hybrid Architectural Approach to Understanding and Appropriately Generating Indirect Speech Acts

**Gordon Briggs** and **Matthias Scheutz**
Human-Robot Interaction Laboratory
Tufts University, Medford, MA 02155, USA
{gbriggs,mscheutz}@cs.tufts.edu

**Abstract**

Current approaches to handling indirect speech acts (ISAs) do not account for their sociolinguistic underpinnings (i.e., politeness strategies). Deeper understanding and appropriate generation of indirect acts will require mechanisms that integrate natural language (NL) understanding and generation with social information about agent roles and obligations, which we introduce in this paper. Additionally, we tackle the problem of understanding and handling indirect answers that take the form of either speech acts or physical actions, which requires an inferential, plan-reasoning approach. In order to enable artificial agents to handle an even wider-variety of ISAs, we present a hybrid approach, utilizing both the idiomatic and inferential strategies. We then demonstrate our system successfully generating indirect requests and handling indirect answers, and discuss avenues of future research.

## Introduction

*Indirect speech acts* (ISAs) comprise a frequently employed, diverse set of human communication modalities (Searle 1975; Hassell, Beecham, and Christensen 1996). As a result, artificial agents like robots who communicate with humans in natural language dialogues will have to handle typical ISAs, especially if the goal is for the dialogue interactions to be *natural* and intuitive for human interactants (Scheutz et al. 2007). While natural and easy for humans, processing ISAs in artificial agents involves a complex network of different interacting components in a cognitive system.

Two distinct methods of handling ISAs have been identified and developed. The *idiomatic* approach leverages the fact that certain utterances have become conventionally synonymous with other speech acts (e.g., "Can you get me a coffee?" as a request for coffee) (Wilske and Kruijff 2006). The *inferential* approach, on the other hand, uses reasoning based on a mental model of the interlocutor to infer the intended meaning of an utterance (e.g., see (Perrault and Allen 1980) for a plan-based inferential approach). While it is an appealing strategy that in principle can handle both conventionalized and unconventionalized instances of ISAs, it can be computationally expensive and thus may not make sense for conventionalized ISAs.However, for unconventionalized ISAs, these reasoning approaches may need to be implemented.

The contribution of this paper is two-fold. First, we present a hybrid system that incorporates both an idiomatic approach and a plan-based inferential approach to efficiently understand both indirect requests and answers. Second, we show how our system utilizes knowledge of *social roles* and *social norms* to appropriately generate both conventionalized and unconventionalized request ISAs. Previous work in human-robot interaction (HRI) has not explicitly tackled sociolinguistic aspects of ISAs, which is necessary for socially appropriate request formulation. We then proceed to validate both generative and understanding portions of our integrated system through proof-of-concept demonstrations. Finally, future directions for investigating technical challenges and the sociolinguistic aspects of ISAs in HRI are presented.

## Indirect Requests and Answers

Consider a scenario in which a cleaning robot R is tasked by a supervisor (Alice) to sweep the floor of large room in conjunction with another co-worker (Bob). R needs to get Bob to assist with the activity through a verbal request, which can be phrased in direct and indirect ways: "Sweep the floor!", "Can you sweep the floor?", "I need you to sweep the floor.", and "Alice wants the floor swept." However, whether or not these request forms are considered socially appropriate depends on the mutually understood social relationship between R and Bob as well as the interaction context.

Brown and Levinson (1987) articulate conversational strategies used to convey varying degrees of politeness during natural language interaction. The strategies relevant to formulating requests are summarized below, in general order of least to most polite:

1. Be direct (bald on-record).
2. Be conventionally indirect.
3. Be unconventionally indirect (off-record).
4. Refrain from the request.

In addition to considerations of politeness, the appropriate request strategy is dependent on knowledge of social roles and norms (Locher and Watts 2005; Allwood, Traum, and Jokinen 2000). Table 1 describes the appropriateness of request strategies in the room sweeping scenario depending on social relationship and interaction context. If a speaker

| Request Utterance | Social Role | Politeness Norm | Appropriateness |
|---|---|---|---|
| Direct Request: (e.g. "Sweep the floor") | Manager | Politeness Favored | Non-politic (Too Direct - FTA) |
| | | Directness Favored | Politic |
| | Co-worker | Politeness Favored | Non-politic (FTA) |
| | | Directness Favored | Non-politic (FTA) |
| Conventionally Indirect Request: (e.g. "Can you sweep the floor?", "I would like you to sweep the floor.") | Manager | Politeness Favored | Politic |
| | | Directness Favored | Non-politic (Too Indirect) |
| | Co-worker | Politeness Favored | Non-politic (FTA) |
| | | Directness Favored | Non-politic (FTA) |
| Unconventionally Indirect Request: (e.g. "Alice wants the floor swept.") | Manager | Politeness Favored | Politic |
| | | Directness Favored | Non-politic (Too Indirect) |
| | Co-worker | Politeness Favored | Politic |
| | | Directness Favored | Politic |

Table 1: Appropriateness of certain speech acts in a given social context. Speech acts can be considered inappropriate (non-politic) for implying excessive or wholly unsupportable social dominance over an interlocutor (FTA) or for being too indirect in contexts where directness is desired.

is in a peer relationship with an interlocutor, the interlocutor may not be obligated to acquiesce to requests to perform actions on behalf of the speaker. If such a request were made (whether direct or conventionally indirect), a socially undesirable *face-threatening act* (FTA) would occur. An alternate way of getting the interlocutor to adopt the desired goal would be an unconventionally indirect approach: informing Bob of the desire of his supervisor will trigger a social obligation to fulfill that goal.

This unconventionally indirect approach is not necessary, however, in the case where the requester is in a supervisory position over the interlocutor, such that the interlocutor does have an obligation to obey a command. Yet, in this case, socially appropriate utterance selection is still modulated by social knowledge about the expected level of directness in the interaction context. For instance, certain interaction contexts require direct language to avoid confusion (e.g., military or medical settings), whereas in other contexts politeness is favored. Therefore, it is necessary for R to have both knowledge about social roles and obligations as well as social norms of directness within the interaction context in order to select the appropriate request strategy.

Now we will examine the issue of indirect answer handling. Consider another scenario in which R has a goal to sweep the floor of the mailroom. Here, R does not know where the mailroom is located, but generates a plan to acquire this information through a dialogue interaction.

Example 1:
Robot: "Do you know where the mailroom is?"
Bob: "It is the second door on the right."

Although Bob has not directly answered the question, it is clear that he has correctly understood the intention of R, interpreting the question as an indirect request for the location of the mailroom.

Example 2:
Robot: "Do you know where the mailroom is?"
Bob: "Follow me."

Bob has again not directly answered the question, but instead issued a command to R. However, it would be clear to a human that this speech act is likely part of an *alternate plan* to help R achieve what Bob believes is R's goal. R ought to be able to ascertain whether or not such an indirect response is indicative of goal comprehension and adapt its behavior accordingly (either adopting the alternate plan or politely rejecting it, e.g. "Thank you for the offer, but I don't want to go there right away.").

Example 3:
Robot: "Do you know where the mailroom is?"
Bob: [gestures and walks off]

Although Bob does not respond with a speech act, the observed action is possibly relevant and again indicative of an alternate plan to help the robot achieve the goal. Not only will the robotic agent require the ability to interpret a variety of speech acts, but also the ability to make sense of non-linguistic actions in the context of a task-based dialogue.

## Proposed Hybrid Approach

We first begin with an architectural overview, followed by a discussion of how we handle conventionalized indirect requests. We then proceed to present our approach to handling unconventionalized ISAs (with a focus on indirect answers). Finally, we discuss how socially sensitive natural language generation is performed.

The architectural diagram of the various involved components in Figure 1 shows two distinct input pipelines through which the robot gathers data from the external world: the NL pipeline and the non-linguistic perceptual pipeline. Components in the system architecture were developed in the Agent Development Environment (ADE) (see http://ade.sourceforge.net/) which is a framework for implementing distributed cognitive robotic architectures.

*The NL pipeline*: The automatic speech recognition system (ASR) sends detected natural language text to the NLP component, which performs parsing, reference resolution, and initial semantic analysis. The results are sent to the pragmatics component which makes pragmatic adjustments and passes the final semantic analysis to the robot's belief sys-

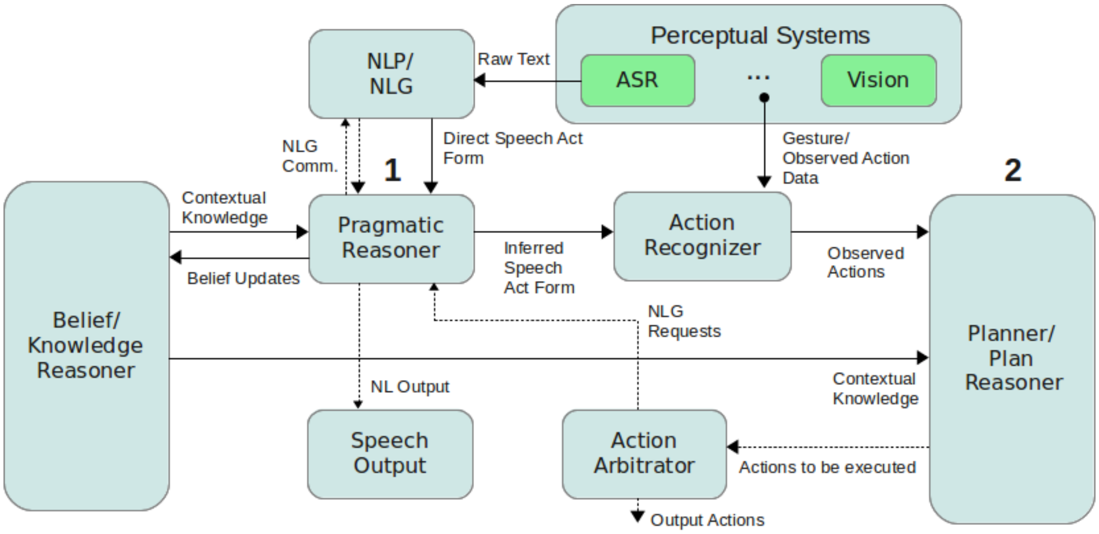


Figure 1: A diagram of the robotic architecture and component interactions involved in ISA understanding and generation. Note that the component indicated by the (1) implements the *idiomatic* understanding approach, while the component indicated by the (2) implements the plan-based *inferential* approach. The dotted lines indicate data flow for action and NL output, whereas the solid lines indicate data flow for action and NL understanding.

tem (specifying how the robot should update its own beliefs about the world, as well beliefs about other agents and their beliefs). Updates are also sent to the action recognizer component that maps the observed NL input, if possible, to the corresponding speech action understood by the robot's plan-reasoning system.

*The non-linguistic pipeline*: A variety of perceptual systems (e.g., the robot's vision system) detect data that is relevant to ascertaining the type of non-linguistic action the interactant is performing. This data is forwarded to the action recognizer component for action identification purposes.

Handling of conventionalized, or idiomatic, ISAs is achieved in the NL pipeline, as the idiomatic approach leverages only linguistic rules. A brief discussion of the pragmatic reasoner that will handle these ISA forms is found in the subsequent section.

## Handling Idiomatic ISAs

The pragmatics component (1) has several roles. It is responsible for determining updates to the agent's own beliefs based on received input from the NLP component as well as handling NLG requests from other components (e.g., planner). The pragmatic reasoning component receives surface semantics and utterance information from the NLP component in the form $u = UtteranceType(\alpha, \beta, X, M)$, in which $UtteranceType$ denotes the utterance classification, $\alpha$ denotes the speaker identity, $\beta$ denotes the addressee identity, $X$ denotes the surface semantics of the utterance, and $M$ denotes a set of sentential modifiers. The pragmatics component contains a set of pragmatic rules $\Sigma = \{\sigma_1, ..., \sigma_n\}$, in which rule $\sigma$ has a set of contextual constraints $C$, an utterance form $U$, and a set of belief updates $[[U]]_C$. These rules, for instance, can be used to infer the belief-model implications of various adverbial modifiers (Briggs and Scheutz 2011).

The functional role of the pragmatics reasoner is to determine the underlying, intended semantics of an utterance given the surface semantics and the context. This makes it the ideal component to implement the idiomatic approach to handling conventionalized ISAs, which involves interpreting speech acts of one form as speech acts of another form in certain contexts. However, when such linguistic conventions are not available, deeper reasoning is required. The components that carry out the higher level inferential reasoning involved in tackling unconventionalized ISAs are introduced in the following section.

## Handling Unconventionalized ISAs

As illustrated in Figure 1, the core of our inferential ISA reasoning system is contained in a planner/plan-reasoning system. This component receives state information from the robot's belief/knowledge storage and reasoning component as well as information regarding observed speech acts and non-linguistic actions from an action identification component. We first describe the planner implementation, followed by the various sense-making plan-reasoning algorithms that are employed by the reasoning component.

**The Planner.** The planner/plan-reasoner contains STRIPS-style action definitions for a variety of speech acts (including those defined in Perrault and Allen (1980)) and non-linguistic actions. Planning is implemented via the GraphPlan algorithm (Blum and Furst 1995). In the following section, we describe the mechanism used to handle unexpected actions, in particular indirect answers, in the course of a robot's plan-execution.

**Sense-making through Plan-reasoning.** Let us consider a scenario that involves the robot $R$ attempting to achieve a goal $g$. Let $\Phi$ denote the set of all valid plans generated by the planner to achieve $g$, and let $\phi_0 \in \Phi$ denote the plan currently adopted by $R$ to achieve the goal. Let $\phi$ denote a plan consisting of a sequence of actions $[a_1, ..., a_n]$. The robot tracks its progress in $\phi_0$ with a counter value $c$, such that $\phi_0(c)$ is the current action being executed either by the robot or by another agent it is interacting with. When $\phi_0(c)$ is an action the robot is not performing, the robot must wait until an action $a_{obs}$ is observed and identified by the action recognition component. The planner/plan-reasoning system handles observed actions based on the following cases:

- (Case 1) $a_{obs} = \phi_0(c)$. In this case, the plan is proceeding as expected. This corresponds to a direct answer (or a conventionally indirect answer).
- (Case 2) $a_{obs} = \phi_0(c')$, $c < c'$. In this case, an action is observed that is consistent with the current plan, but involves a possible jump ahead in the plan. $c$ should be set to $c'$ if such a jump still results in desired goal state. If this check fails, alternate plans should be considered (case 3).
- (Case 3) $\neg\exists c'$ such that $\phi_0(c') = a_{obs}$ but $\exists \phi \in \Phi, \phi \neq \phi_0$ such that $a_{obs} \in \phi$. In this case, the observed action is inconsistent with the current plan, but it can be found in an alternate plan that has been calculated. In this case the alternate plan $\phi$ should be adopted as the current plan $\phi_0$ and the counter should be set such that $\phi(c) = a_{obs}$, pending a validity check as described in the previous case.
- (Case 4) $\neg\exists \phi \in \Phi$ such that $a_{obs} \in \phi$. In this case, the observed action cannot be accounted for by any known plan. It is likely that the robot's interaction partner has misunderstood the robot's goal (or the robot does not have enough knowledge to have discovered the alternate plan of the other agent).

This sense-making algorithm, which utilizes the set $\Phi$ of plans generated by the robot to achieve a goal, relies on an assumption that the robot's interactant has correctly understood the robot's intentions. We justify such an assumption through the Gricean Maxim of Relation, which states that any contribution to a cooperative interaction must "be relevant" (Grice 1975). However, in case 4, in which no known plan provides an explanation for the interactant's action, the robot ought to conclude that such an assumption was in error. In this case the robot either must initiate an abductive-reasoning process to infer what alternative goal its interactant has ascribed to it and/or explicitly communicate its intentions (e.g., "I'm sorry, I don't understand, I am trying to get to mailroom.").

**Modeling Social Roles.** Many actions include as a precondition that the agent must want to perform that action. To explicitly model agents deciding to adopt another agent's goal as their own (in response to a request, for instance), Perrault and Allen (1980) present a "glue" action $cause\_to\_want$:

action: $cause\_to\_want(\alpha, \beta, g)$
type: $agent(\alpha) \wedge agent(\beta)$
precond: $B(\beta, W(\alpha, g))$
effect: $W(\beta, g)$

The $cause\_to\_want$ action assumes that agent $\beta$ will potentially adopt any goal from any agent as its own, which is unrealistic. Though the process by which an agent decides to adopt another agent's goal can be quite complex, social obligations can be used as one straightforward way of defining certain instances where such goal adoption occurs. For instance, in our cleaning agent example, we can argue that there exists an obligation by an agent $\beta$ to want to sweep a location $L$ if $\beta$ believes that his or her supervisor $\alpha$ wants this to be accomplished and he or she has a janitorial role. Below we define this obligation glue action:

action: $obl\_sweep(\alpha, \beta, L)$
type: $agent(\alpha) \wedge agent(\beta) \wedge location(L)$
precond: $B(\beta, role(\beta, janitor)) \wedge$
$B(\beta, is\_supervisor\_of(\alpha, \beta)) \wedge$
$B(\beta, W(\alpha, swept(L)))$
effect: $W(\beta, sweep(\beta, L))$

By replacing the general $cause\_to\_want$ glue action with a series of obligation modeling actions, socially sensitive dialogue interaction planning can be undertaken. Socially inappropriate speech acts will not be utilized in communicative plans, because they will not achieve the desired effects to the interlocutor's belief state. A walkthrough of the generation process will be shown in greater detail later.

## Generating Requests

When a robot desires to get its interlocutor $\beta$ to adopt a goal $g$, a two-step process is initiated where first the appropriate speech acts must be selected (what to say), followed by a step that determines what utterances to use to carry out the selected speech acts (how to say it).

**What to Say.** The planner is initialized in order to formulate a dialogue plan that will result in the interlocutor adopting the goal $g$. As described previously, this dialogue planning process is what determines whether or not the request can be made via a direct statement of the speaker's desire (either through bald on-record directness or conventional indirectness) or must be implied by an unconventionally indirect utterance. It may be the case that no plan exists to get $\beta$ to adopt goal $g$, meaning that given the social constraints, such a request is not possible.

**How to Say It.** Once a valid dialogue plan has been found, the constituent speech act(s) are sent to the pragmatics component. A set of applicable pragmatic rules that convey the appropriate propositions are found. For the purposes of assisting NL generation, each pragmatic rule $\sigma$ has an associated value $\eta$, which denotes the degree to which the utterance can be construed as a face-threatening act (Gupta, Walker, and Romano 2007), as well as a marker denoting whether or not it is a direct-form (i.e., whether inferred semantics are the same as the semantics of the surface form). Currently, $\eta$ values are assigned according to the surface utterance type associated with the pragmatic rule. Direct commands (e.g., "Sweep the floor") are given a high $\eta$ level, whereas questions and statements are given a low $\eta$ value.

The aim of the utterance selection algorithm is to be as direct as possible, while not exceeding the maximum allowable $\eta$ level, which is determined by social context. For instance, a workplace environment that requires directness in communication would prompt a higher maximum $\eta$ level to be set. $\eta$-levels can also be determined by social power relationships, as superiors will have higher allowable $\eta$ levels when communicating with subordinates than vice versa. This $\eta$-level mechanism is what generates the use of conventional ISA forms. An $\eta$ level limit may prohibit the use of a direct command, but if the same semantic content can be communicated using an ISA (being either a statement or a question), than this indirect form is used for the request.

## Request Generation Walkthrough

We revisit the example from the beginning of the paper, in which a robot is attempting to get a co-worker to assist it in sweeping the floor. Denote the robot as agent $R$, the co-worker Bob as $\beta$, the supervisor Alice as $\alpha$, and the floor as $L$. We examine three cases: one in which the robot has a peer relationship with $\beta$; one in which the robot has a supervisory relationship with $\beta$ and the work relationship has an expectation of politeness; and one in which the robot has a supervisory relationship with $\beta$ and the work relationship has an expectation of directness.

**Peer Relationship.** In this case, the obligation modeling action $obl\_sweep(R, \beta, L)$ cannot be triggered as the precondition $B(\beta, is\_supervisor\_of(R, \beta))$ does not hold. However, $obl\_sweep(\alpha, \beta, L)$ can be triggered. Therefore, the planner finds the plan beginning with, $inform(R, \beta, want(\alpha, swept(L)))$. Because this speech action is conveyed by a statement (which has a low associated $\eta$ value), the contextual preference for direct vs. polite speech does not affect the utterance selection process. The pragmatics component selects the utterance form $Stmt(R, \beta, W(\alpha, swept(L)), \{\})$ to convey. Finally, the NLG component translates this utterance to the appropriate NL sentence: “Alice wants the floor swept.”

**Supervisory Relationship - Low Directness.** In this case, the obligation modeling action $obl\_sweep(R, \beta, L)$ can be triggered. Therefore, the planner finds the plan beginning with the direct request speech act $request\_sweep(R, \beta, L)$. However, in this case, utterances with high $\eta$ levels cannot be used. As such, a direct command utterance of the form $Instruct(R, \beta, sweep(\beta, L), \{\})$ cannot be selected. Instead, the conventionally indirect form $AskYN(R, \beta, can\_do(\beta, sweep(\beta, L)), \{\})$ can be selected. Thus, the NLG component translates this utterance form to the NL sentence: “Can you sweep the floor?”

**Supervisory Relationship - High Directness.** As in the previous case, the planner finds the plan beginning with $request\_sweep(R, \beta, L)$. In this case, utterances with high $\eta$ levels can be used. As such, the direct command $Instruct(R, \beta, sweep(\beta, L), \{\})$ is valid (and favored over more indirect utterances). Thus, the NLG component translates this utterance form to the NL sentence: “Sweep the floor.” Table 2 summarizes these results.

## Indirect Answer Understanding Walkthrough

We first step through the operation of the indirect answer handling system and the simulation environment used to test our approach. Then we give examples of how the system can find a variety of alternate plans to make sense of indirect answers and discuss examples that we have tested in simulation.

### Simulation Environment

ADE provides a simulator that simulates the physical environment and various robotic platforms in real-time. Hence, changes needed to adapt system components for use in the simulation environment to use in the real-world are minimal. Regardless, all non-plan-reasoning components have previously been deployed and verified on physical robots.

We initialized the simulation environment map to reflect the area around our laboratory. Two simulated agents were instantiated: one simulated Videre Era robot, which was backed by a streamlined version of the architecture in Figure 1[1], and one simulated Adept Pioneer robot (which we designated as the simulated “human” interactant “Bob”). For the simulations, we revisited the scenario in which a service robot must get to a building’s mailroom. In this case, we gave the Videre robot the goal to be at the mailroom. The two robots were placed near each other in a hallway and were aware of each other’s presence. The simulated interactant was programmed to respond in a couple different ways to duplicate some of the exchanges discussed previously (and will be illustrated below).

### Plan Discovery

The Videre robot knows of a set of locations $L = \{bathroom, cafeteria, mailroom\}$ (but not their locations within the building) and an online building map $internet\_map$. The robot spots a nearby “human”, Bob, which it adds to the set of possible interactant agents $A = \{self, bob\}$. The robot has a goal to be at the mailroom, which is represented by the predicate, $want(self, at(self, mailroom))$.

Figure 2 illustrates a small set of possible plans the planner has discovered that would enable the robot to achieve its goal to be at the mailroom. Because the only plan in this set of plans that begins with a self-initiated action is P1, the robot adopts that plan, though it would usually be the case that the robot must choose from a set of self-initiative plans based on some cost metric.

The plans found in the figure correspond to a variety of possible interaction exchanges. For instance, P1 would make sense of example 1. P2 would make sense of the following exchange:

Robot: “Bob, do you know where the mailroom is?”
Bob: “I’ll look that up online...”

P4 would make sense of the following dialogue:

Robot: “Bob, do you know where the mailroom is?”
Bob: “I think you can look that up online...”

[1]The full NL system was not run, as these simulations were intended to focus on the plan-reasoning components.

| Social Role | Directness Level | Planned Speech Act | Utterance Selected |
|---|---|---|---|
| Manager | High Directness | $request\_sweep(R, \beta, L)$ | $Instruct(R, \beta, sweep(\beta, L), \{\})$ |
| | Low Directness | $request\_sweep(R, \beta, L)$ | $AskYN(R, \beta, can(\beta, sweep(\beta, L)), \{\})$ |
| Co-worker | High Directness | $inform(R, \beta, want(\alpha, swept(L)))$ | $Stmt(R, \beta, want(\alpha, swept(L)), \{\})$ |
| | Low Directness | $inform(R, \beta, want(\alpha, swept(L)))$ | $Stmt(R, \beta, want(\alpha, swept(L)), \{\})$ |

Table 2: Speech act and utterance form selected by NL system to get $\beta$ to sweep the floor considering social role and norms.

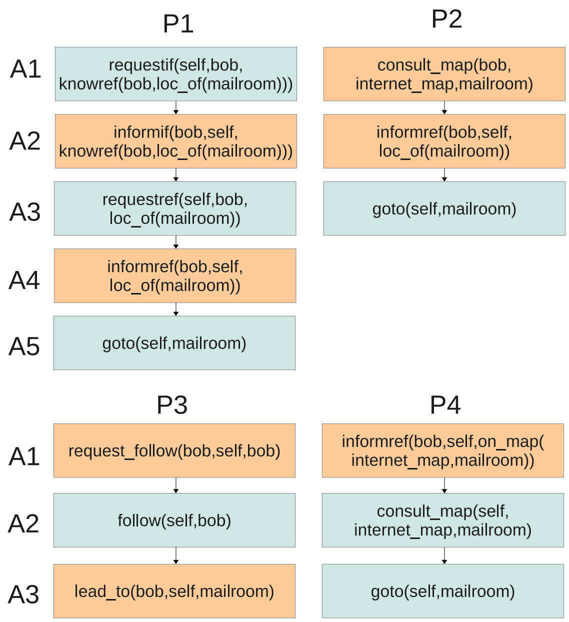


Figure 2: Examples of plans found by the planner to achieve the goal $at(self, mailroom)$. The blue boxes correspond to actions undertaken by the robot, while the orange boxes correspond to actions undertaken by the other agent (Bob).

### Online Plan-Reasoning

We have so far shown how the planning system can discover plans that are capable of making sense of indirect answers by interactants. Below we will examine how the sense-making algorithm handles two example scenarios.

In example 1, the Videre robot has adopted the plan P1 (see Figure 2). The robot executes the $requestif$ action (A1), asking Bob whether or not he knows the location of the mailroom. Instead of directly responding, "yes", Bob responds by providing the robot with the location information of the mailroom, "It is the second door to the right." This utterance corresponds to action A4 in plan P1, and as such falls under Case 2 in the sense-making algorithm. As a result, the current plan-index $c$ is set to A5, and subsequently the robot executes A5 to navigate to the mailroom.

In example 2, the Videre robot has adopted plan P1. The robot again executes action A1, asking Bob whether or not he knows the location of the mailroom. In this case, instead of directly responding, Bob responds with the command "Follow me." The NL system determines this utterance corresponds with the speech act, $request\_follow(bob, self, bob)$, which is not found in the current plan P1. However, it discovers that this action corresponds with action A1 in plan P3, so Case 3 in the sense-making algorithm applies. As a result, the current plan is set to P3, and the current plan-index $c$ is set to A2. Subsequently, the robot executes A2 and begins to follow Bob.

## Future Work

While the NL system is able to identify speech acts for the plan-reasoner, other mechanisms are needed for non-linguistic action recognition. Previous approaches to this problem have included using Hidden-Markov Models (HMM) to identify movement activities based on lower-level perceptual data (Kelley et al. 2008; 2012). A series of such activity recognizers will be required to enable detection of all the non-linguistic actions the plan-reasoner has knowledge of. High-level knowledge of likely activities from the plan-reasoning component could constrain which detectors are active and bias the decisions of the operational detectors.

It has been assumed based on general human-computer interaction research that humans will fallback on human-human social behaviors when interacting with computers (Nass, Steuer, and Tauber 1994), and as such will deploy ISAs with robots (Wilske and Kruijff 2006). However, as previously mentioned, ISAs may not be necessary/present in highly task-oriented dialogues between well-defined superiors and subordinates (which humans and robots will likely be, respectively). To clarify this issue, we will need to perform HRI experiments to test this hypothesis in a variety of HRI contexts.

## Conclusion

The ability for dialogue agents to phrase utterances in socially appropriate ways requires that knowledge of social roles, obligations, and norms be utilized throughout the cognitive architecture. We presented a hybrid approach integrated into a cognitive architecture that implements both the idiomatic and inferential approaches to ISA handling and shown how this social knowledge can be integrated in the system to achieve socially sensitive NL generation. In addition, we showed how this system can make sense of indirect answers, which can take the form of both speech acts and physical actions. Finally, we presented a proof-of-concept validation for this integrated hybrid approach for both socially sensitive utterance selection and indirect answer understanding.

## Acknowledgments


This work was supported in part by ONR grant #N00014-11-1-0493 and in part by NSF grant #111323.

## References


Allwood, J.; Traum, D.; and Jokinen, K. 2000. Cooperation, dialogue and ethics. *International Journal of Human-Computer Studies* 53(6):871–914.

Blum, A. L., and Furst, M. L. 1995. Fast planning through planning graph analysis. *ARTIFICIAL INTELLIGENCE* 90(1):1636–1642.

Briggs, G., and Scheutz, M. 2011. Facilitating mental modeling in collaborative human-robot interaction through adverbial cues. In *Proceedings of the SIGDIAL 2011 Conference*, 239–247. Portland, Oregon: Association for Computational Linguistics.

Grice, P. 1975. Logic and conversation. In Cole, P., and Morgan, J., eds., *Syntax and Semnatics, 3: Speech Acts*.

Gupta, S.; Walker, M. A.; and Romano, D. M. 2007. How rude are you?: Evaluating politeness and affect in interaction. In *Affective Computing and Intelligent Interaction*. Springer. 203–217.

Hassell, L.; Beecham, S.; and Christensen, M. 1996. Indirect speech acts and their use in three channels of communication. In *In Proceedings of the First International Workshop on Communication Modeling - The Language/Action Perspective*.

Kelley, R.; Tavakkoli, A.; King, C.; Nicolescu, M.; Nicolescu, M.; and Bebis, G. 2008. Understanding human intentions via hidden markov models in autonomous mobile robots. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, HRI '08, 367–374.

Kelley, R.; Wigand, L.; Hamilton, B.; Browne, K.; Nicolescu, M.; and Nicolescu, M. 2012. Deep networks for predicting human intent with respect to objects. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, HRI '12, 171–172. New York, NY, USA: ACM.

Locher, M., and Watts, R. 2005. Politeness theory and relational work. *Journal of Politeness Research. Language, Behaviour, Culture* 1(1):9–33.

Nass, C.; Steuer, J.; and Tauber, E. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, 72–78. ACM.

Perrault, C. R., and Allen, J. F. 1980. A plan-based analysis of indirect speech acts. *Comput. Linguist.* 6(3-4):167–182.

Scheutz, M.; Schermerhorn, P.; Kramer, J.; and Anderson, D. 2007. First steps toward natural human-like HRI. *Autonomous Robots* 22(4):411–423.

Searle, J. 1975. Indirect speech acts. In Cole, P., and Morgan, J., eds., *Syntax and Semnatics, 3: Speech Acts*.

Wilske, S., and Kruijff, G.-J. M. 2006. Service robots dealing with indirect speech acts. In *IROS'06*, 4698–4703.