A Cognitive Framework for Delegation to an Assistive User Agent

Karen Myers and **Neil Yorke-Smith**
Artificial Intelligence Center, SRI International, Menlo Park, CA 94025, USA
{myers,nysmith}@ai.sri.com

Abstract

We present a BDI-based framework for a cognitive agent that acts as an assistant to a human user by performing tasks on her behalf. This framework accommodates: (1) user delegation of tasks to the agent, which may be inconsistent or infeasible or both, and (2) user-specified guidance and preferences on the execution and management of these tasks, which we call *advice*. Within this *delegative BDI* model, we propose a combined mechanism for the problems of goal adoption and commitment revision, with the aim of a well-founded, practical framework for an assistive agent with extended life-cycle.

Introduction

Our interest is to develop intelligent agents that can assist a human by performing and managing tasks on her behalf. These interests are being pursued in a project whose aim is to build an intelligent personal assistant called *CALO* (Cognitive Assistant that Learns and Organizes) (SRI International 2005) that will support a high-level knowledge worker such as a lab director or senior manager. CALO will be able to perform routine office tasks for its user (e.g., arrange meetings, complete online forms, file email), manage open-ended processes (e.g., purchase a computer, arrange a conference), and anticipate and act on future needs of its user.

In this paper, we propose a cognitive framework for an intelligent assistant, inspired by our work on CALO. This framework is intended to support a *delegative* model of intelligent assistance, in which the user assigns tasks to the automated assistant with the expectation that it will complete them in a mostly autonomous manner. Thus, the agent may interact with the user to clarify her desires or to solicit inputs to assist with problem solving at various stages, but the agent has primary responsibility for completing the assigned tasks. This approach is to be contrasted with a *collaborative* model, in which the agent and human work together to complete a shared task (Allen, Blaylock, & Ferguson 2002).

Our design builds on the Belief-Desire-Intention (BDI) model of agency (Rao & Georgeff 1991), which has become the predominant architecture for the design of cognitive agents. The BDI model provides an explicit, declarative representation of three key mental structures of an agent: informational attitudes about the world (beliefs), motivational attitudes on what to do (desires), and deliberative commitments to act (intentions). This explicit representation enables ready inspection of an agent's operation, both by an observer and by the agent itself, thus supporting reflective capabilities such as explanation and redirectability.

Two essential characteristics underpin the BDI model. The first is the distinction between an agent's desires, which may be conflicting or even impossible, and the *goals* that it chooses to pursue. An agent will generally not act on all of its desires, but rather will employ some mechanism to select from among its motivations a subset to address. The second characteristic is the ability for an agent to commit to activities via its intentions. Thus, rather than having to reevaluate its motivations at each step, an agent can commit to a collection of activities, reconsidering them only when world dynamics have changed sufficiently. Thus, intentions provide the means for an agent to focus its attention on the pursuit of selected goals.

The BDI model provides a theory of agent operations that is neutral with respect to the *type* of an agent. Assistive agents such as CALO constitute one such type. Others include self-interested agents whose sole objective is to achieve their own desires (Rao & Georgeff 1991), and cooperative agents that work within a team (Ancona & Mascardi 2003). Part of the contribution of this paper is an augmentation of the basic BDI model to reflect the specialized cognitive and interaction requirements of assistive agents.

Our proposed framework also addresses certain recognized gaps within the general BDI theory (e.g. (Rao & Georgeff 1995; Winikoff *et al.* 2002; Ancona & Mascardi 2003; Dastani & van der Torre 2004; Morley & Myers 2004)). These gaps derive from the fact that most BDI model work has focused on conceptual rather than actual agents. In particular, there has been insufficient regard to the embodiment of the theory within agents that operate in the real world. As a result, certain critical mechanisms required for a functional BDI agent are not well defined. So, for example, what principles should an agent use in identifying a subset of its motivations upon which to act (Dastani & van der Torre 2004)? Under what conditions should an agent reconsider its commitments (Schut, Wooldridge, & Parsons 2004)? At the same time, the bulk of BDI implementations has lacked formal grounding, in part because of pragmatic needs and in part because of the gaps in the theory. For example, most implemented systems have at best an implicit representation of goals, com-

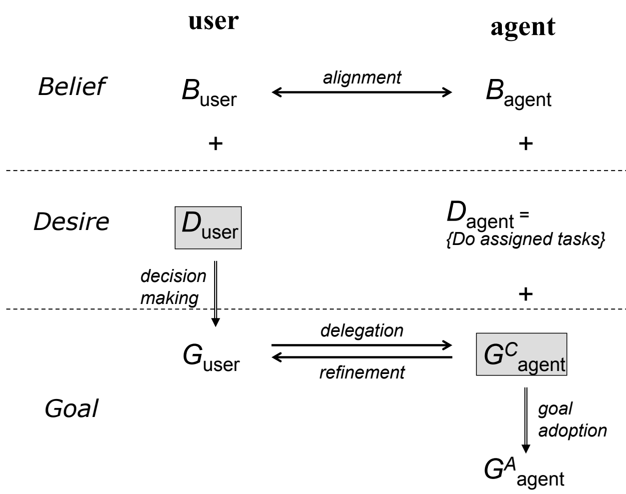

Figure 1: User and Agent Cognitive Models. The agent's sole desire is to satisfy the goals delegated to it by the user. To do so, it must maintain a consistent and feasible set of adopted goals to execute. That may require interactions with the user to align differing beliefs and to refine assigned tasks. D_{user} and G^C_{agent}, which are shaded in the figure, may be inconsistent with the beliefs of the user and agent respectively.

bining desires and intentions into a single deliberative execution model (Thangarajah, Padgham, & Harland 2002; Ancona & Mascardi 2003). A second problem is that agent models and systems to date have focused on a short-term view of agent execution, rather than considering persistent agents with long-term extent. One exception, not based on a BDI model, is ICARUS (Choi *et al.* 2004).

To summarize, the key features of our model of delegative assistance are: (1) user delegation of tasks to the agent, possibly involving inconsistent or infeasible tasks, which form the *candidate* goals of an assistive agent; and (2) user-specified guidance and preferences on the execution of these tasks, and on the agent's cognition, called *advice*. Within this model, we propose a combined mechanism for the problems of goal adoption and commitment revision, with the aim of providing a well-founded, practical framework for an assistive agent such as CALO.

Extending the BDI Framework

The sole function of an agent acting as a human's assistant is to perform tasks delegated to it by the user. As such, the agent has no personal desires that drive its behavior. For this reason, an assistive agent can be viewed as having the single, fixed desire of satisfying the user's taskings.

The user applies some form of decision-making process (its nature is not relevant to this paper) to identify tasks to be assigned to the agent. As is well known from theories in economics and psychology, human decision making is not always rational (Levi 1997), and in any case may be based on limited or outdated information. As such, the tasks given to the agent may be inconsistent with one another, the state of the world, and the means of the agent to act upon them.

An assistive agent may have more knowledge than its user as to task feasibility or appropriateness in certain situations, and can provide value by identifying such problems for the user. Problem identification should initiate a cycle of interactions between the human and agent to revise problematic taskings before commitments to act are made. Because the problems identified by the agent may derive from a misalignment between the agent and human's beliefs about the world, this refinement process may also require the establishment of mutual belief about relevant world conditions.

The potential for infeasible taskings introduces the need for a richer model of motivational attitudes for an agent than the standard BDI theory provides. To this end, we distinguish between motivational attitudes that may be not consistent and feasible (*candidate goals*, we will call them) and those that must be (*goals*, we will call them). As discussed by Dignum, Kinny, & Sonenberg (2002) and others, many formalizations of BDI equate these two types of goals.

The user and the agent are equal neither in authority nor in cognition. In contrast to Perrault & Allen (1980), we do not consider the agent to be 'human-like' in its nature. Even so, Figure 1 shows how there is a parallel in the reasoning process that the human and agent undertake. The user holds that her beliefs are true, to the limit of her knowledge. She has desires that may or may not be mutually consistent and feasible. She decides on goals that are a consequence (by some 'rational' process (Levi 1997)) of her beliefs and desires. The agent also holds that its beliefs are grounded in its world knowledge. Its candidate goals (G^C) are identical to the user's goals; the candidate goals may or may not be mutually consistent and feasible relative to the agent's beliefs. By application of a rational reasoning process, which may involve interactions with the user that lead to modifications of her goals and beliefs or the agent's beliefs, the agent derives an appropriate set of adopted goals (G^A) upon which it will act.

An agent such as CALO persists, collaborates, and learns over an extended period of time. To support operation throughout this extended life-cycle, a further extension to standard BDI models of goals is necessary. For instance, the user may task the agent with a candidate goal that is not possible to achieve at this time (purchase a book), but will be possible at some point in the future (when new stock arrives with the seller). Thus, the set of candidate goals comprises newly arriving tasks from the user, together with previously unadopted — but not yet expired — previous taskings.

The separation of feasible and infeasible goals and support for persistence of goals constitute the first extension to the BDI framework that we propose. The second extension, which arises from the nature and role of an assistive agent, is to incorporate a notion of *advisability* of an agent by a user. Advisability will enable the user to direct an agent on how to accomplish assigned tasks, and on the degree to which the agent can make decisions without human intervention (so-called *adjustable autonomy* (Maheswaran *et al.* 2003)). Some implemented BDI systems, such as SPARK, include *execution advice* of this type (Myers & Morley 2003). The assistive framework presented in this paper extends execu-

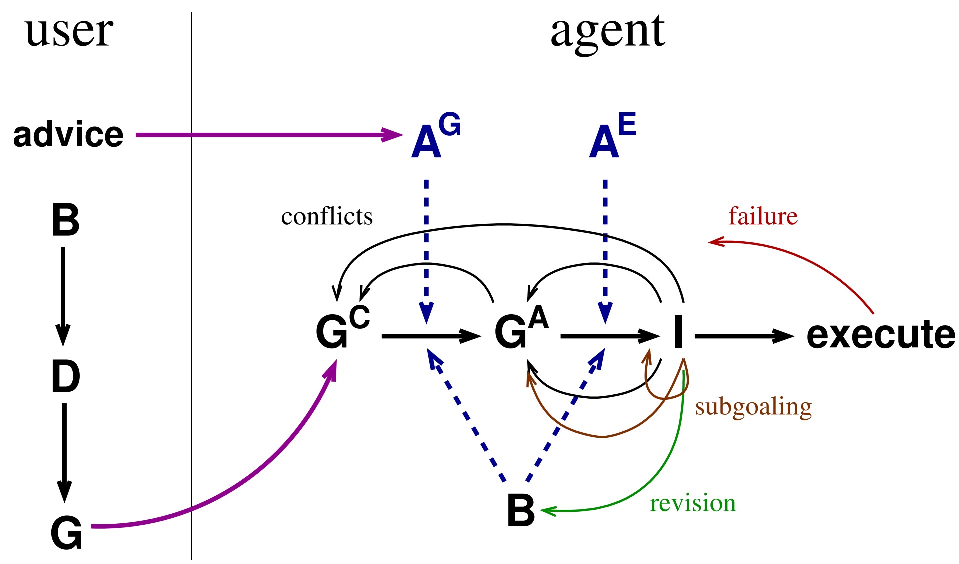

Figure 2: The delegative BDI agent architecture

tion advice to include guidance on the process of adopting and dropping goals, in particular, how to respond to problematic candidate goals. We call this *goal advice*. Although not considered in this paper, advisability could further include guidance on how to resolve conflicts among advice (Myers & Morley 2002), and conflicts between advice and other components in the agent's deliberation.

A Delegative BDI Framework

To meet these multiple requirements, we propose an extended BDI framework that includes Advice (A), Desires (D), and candidate and adopted Goals (G^C and G^A). Figure 2 outlines the framework. On the left is a depiction of the user's decision-making. Arising from some model of beliefs and desires — typically, probabilities and utilities, but the nature and model of the user's reasoning are not important to the framework we propose — come the user's decisions.

The delegative agent is depicted on the right side of Figure 2. Besides perception of the world, which updates the agent's beliefs, the input to this model of agency comes from the user in two forms. First and primary, the user specifies tasks to the agent, which become the agent's candidate goals (G^C) as described above (the lower left-to-right arrow). Related, the user specifies advice to the agent (the upper left-to-right arrow), as described below.

Specifically, from its candidate goals, combined with its beliefs, the agent derives adopted goals ($G^C \rightarrow G^A$) by a process called *goal adoption*. From its adopted goals, the agent derives intentions ($G^C \rightarrow I$) and commits to them by the process of *intention generation*. Finally, intentions are executed. In systems such as PRS (Georgeff & Ingrand 1989) and SPARK (Morley & Myers 2004), intention execution involves subgoaling, that is, hierarchical expansion of intentions and generation of new goals. Intention execution may also lead to revision of beliefs and goals, and to conflicts between different cognitive attitudes. Execution brings the possibility of failure, which may have similar ramifications. The next section expands on some of these aspects.

Definitions. The proposed framework is composed of eight principal aspects:

- **Beliefs** are the agent's accepted knowledge about the world and its current state (the standard BDI interpretation). Example: "Alice and Bob are the best fits on paper for the open position".
- **Desires** are states that describe the agent's motivations arising from its nature or type (its 'character', one might say). As discussed, for our purposes the agent, as a delegative user assistant, has the single desire of performing all tasks that the user assigns to it.
- **Candidate Goals** are tasks that provide the agent's motivation.[1] That is, candidate goals characterize tasks that the agent would like to accomplish, given its desire to satisfy the user's taskings, but may not be able to because of extenuating factors. Candidate goals need not be consistent with each other, with the state of the world, or with adopted intentions. For example, "I wish to interview Alice at 9 a.m. on Monday", and "I wish to interview Bob at 9 a.m. Monday" constitute conflicting candidate goals.
- **Goals** are a consistent, feasible set of tasks, which are derived from the agent's candidate goals. As noted, implemented BDI systems suppose candidate goals to be consistent. The next section explains what we mean by consistent and feasible. Example: "I want to invite Alice" (and it's possible).

[1] Theoretical BDI frameworks represent desires as states to be accomplished; in contrast, most implemented systems represent goals as tasks to be performed. For consistency with these conventions, we represent desires as states but goals as tasks. State-to-task translation, while reasonably straightforward, lies beyond the scope of this paper.

- **Intentions** are committed goals together with the means to achieve them (plans). Intentions represent the tasks arising from the user that have successfully passed the conditional aspects; as far as it knows, the agent can achieve these tasks, and it has committed to doing so. Example: "I am inviting Alice" (by emailing her three possible dates).
- **Goal-Advice** are constraints on the adoption of candidate goals as goals. Examples: "Don't invite multiple candidates on the same day", "Attend only one conference this summer".
- **Execution-Advice** are constraints for user directability of problem solving. Whereas Goal-Advice advises the agent what to do, Execution-Advice advises the agent how to do it. Examples: "Use email to invite candidates", "Fly on a US carrier".
- **Plans** are the means to achieve intentions (the standard BDI interpretation). Plans are selected from a library of alternative procedures for achieving an intention.

In this paper we consider a sole assistive agent. While recognizing the importance of multiagent interaction and problem solving, and their relevance to a human assistant such as CALO, we do not consider mechanisms for coordination as obligations and norms. One can envision extending the delegative BDI framework to include these aspects, as for instance Dignum et al. (Dignum *et al.* 2000; Dignum, Kinny, & Sonenberg 2002) do for standard BDI.

Commitment

The issue of commitment lies at the heart of BDI architectures: how should an agent balance its desires and goals with its current set of commitments? This balancing problem embraces two fundamental decisions: determining when to adopt a new goal (*goal adoption*) and determining when to terminate an existing intention (*intention reconsideration*).[2]

The agents literature is for the large part silent on these questions (Thangarajah, Padgham, & Harland 2002; Dastani & van der Torre 2002). In practice, most agent systems skirt the goal adoption problem by making the simplifying assumption that desires and goals can be equated, and by not distinguishing candidate and adopted goals over an extended life-cycle. Intention reconsideration then is driven solely by problems encountered as execution proceeds.

We view both activities as components of the larger cognitive function of deliberating over mental states. For this reason, we propose an approach in which goal adoption and intention reconsideration are inextricably linked in a *commitment deliberation* process. We start by articulating a set of requirements on adopted goals, and then present a general proposal for commitment deliberation.

Requirements on Goal Adoption

We define four conditions that a set of adopted goals *G* must satisfy. Here, consistency implies consistency with respect to a fixed background theory of domain constraints.

- **Self-consistency:** G^A must be mutually consistent
- **Coherence:** G^A must be mutually consistent relative to the current beliefs *B*
- **Feasibility:** G^A must be mutually satisfiable relative to current intentions *I* and available plans
- **Reasonableness:** G^A should be mutually reasonable with respect to current *B* and *I*

The requirements of self-consistency and coherence are common to most BDI frameworks (Rao & Georgeff 1998). Feasibility requires that goals be adopted only if they can be achieved. One issue with imposing a feasibility requirement is that in general, determining achievability of goals is undecidable; in practice, feasibility testing requires a heuristic assessment of the likelihood that the desires could be completed (a form of bounded rationality).

The reasonableness requirement plays an important role in assistive agents, where the user may pose desires that are consistent, coherent, and feasible, but not in her best interest. This type of situation can arise because a user lacks awareness of problem-solving history, commitments, or constraints. For example, suppose that the user asks her agent to purchase a laptop computer; the next day, having forgotten about the request issued the prior day, she requests the agent again to purchase a computer. If the user worked in procurement, such a request may seem reasonable. However, for the typical office worker, the purchase of a computer is a relatively rare event. In this case, it would be helpful for the agent to recognize that the request is unreasonable, and check with the user as to whether she really wants the agent to proceed with the purchase. Note the collaborative process: according to the degree of adjustable autonomy, the agent consults the user in this situation; it does not dismiss the user's tasking out of hand.

Many prior formal models have cast goal adoption as a filtering problem, where some maximal subset of stated desires is identified that satisfies designated requirements (such as those above). We argue that this approach is inappropriate in many situations, as it places undue emphasis on current commitments. More generally, the goal adoption process should admit the possibility of modifying beliefs (through acting to change the state of the world), adopted goals, and intentions in order to enable adoption of new goals.

As a concrete example, suppose that the user desires to attend a conference in Europe but lacks sufficient travel funds to do so. To enable attendance, the user could shorten a previously scheduled trip for a different meeting to save costs, or she could cancel the planned purchase of a new laptop. Alternatively, the user could apply for a travel grant from the department to cover the costs of the European trip.

Commitment Deliberation

A typical BDI agent executes a tight control loop for determining the actions that it will perform. Let $S = (B, G^C, G^A, I)$ denote the agent's current mental state at the start of a cycle through this loop. The mental state also includes desires, but we omit D since our agent has a single,

[2] More generally, it should be possible to suspend and reactivate intentions. Suspension and reactivation can be modeled as a form of intention and adoption; hence, we do not dwell on those topics.

fixed desire. The control flow involves identifying modifications to S from the prior cycle, deciding what to do in response to those changes, and then performing an appropriate set of actions.

Commitment deliberation fits in neatly as an initial step within this control loop. Thus, to start a cycle of the control loop, the agent first performs some deliberation to determine whether it needs to revise its commitments in light of the changes during the past cycle. In general, this deliberation can transition the agent's mental state to a new state S'.

Different types of transitions reflect different strategies to commitment. An *expansion transition* involves the agent adopting an additional goal from its current candidate goals, without requiring modifications to its current adopted goals and intentions. A *proactive transition* involves the agent creating a new candidate goal and adopting it in order to enable some current candidate goal to be addressed at a future stage (e.g., applying for a travel grant to enable the possibility of attending the European conference). A *revocation transition* involves the dropping of a goal together with its corresponding intentions, in order to support the adoption of a different goal (e.g., cancelling a planned trip to free up funds as required to attend a conference).

The goals and intentions for a mental state have associated criteria that influence the deliberation process, including:

- **Goals:** User-specified value or utility (time-varying), priority, and deadline; estimated cost of achievement. For adopted goals, there is also the level of commitment.
- **Intentions:** Cost of change (deliberative effort, loss of utility, delay); level of commitment; level of effort so far (e.g. percentage complete); estimated cost to complete; estimated probability of success (Pfeffer 2005); and the value or utility implied by the goal from which the intention arises.

In selecting goals and intentions, as well as plans to execute intentions, the agent seeks to maximize (or minimize) some combination — sophisticated or otherwise — of these criteria. Hence the move from S to S' can be cast as a multi-criteria optimization problem. Note that some criteria are user-specified while others derive from the agent's cognitive processes. A related problem is considered by Dignum, Kinny, and Sonenberg (Dignum, Kinny, & Sonenberg 2002), in adopting goals based on norms, obligations, and desires.

Being executable tasks, goals and intentions in our framework have utilities attached to their achievement. In the general case, it is well known that there is no way to satisfy all the properties one would like in multi-criteria optimization; utility functions on the components cannot be combined to a single, ideal ranking function on the possible S' (Doyle & Wellman 1991). We envision an agent endowed with a simple strategy for selecting an optimal S' (i.e., there may be multiple such options), and with the capability of this strategy to be overridden by more elaborate strategies specified by the user for a given situation. The strategies, drawn from the literature on constrained multi-criteria optimization (Steuer 1986), need not require quantitative utilities, and need not require that the utilities of the different components be put on a single commensurate valuation scale.

Moreover, the possible moves $S \rightarrow S'$ are constrained by the user's goal advice A^G. Such advice defines user preferences over allowed moves, for example, "Refrain from dropping any intention that is more than 70% complete". Since advice is soft, the agent may consider advice-infeasible moves, with a suitable penalty in the overall assessment.

For an assistive agent, the default strategy can benefit from collaboration with the user, according to the level of adjustable autonomy. We envision that an agent like CALO would consult the user if it cannot establish a clear best S'. For example, it might ask: "Should I give up on purchasing a laptop, in order to satisfy your decision to travel to both AAAI and AAMAS?" Further, over its extended life-cycle, CALO would learn and refine its model of the user's preferences according to her responses in such situations; the time management component of the current CALO system already undertakes this type of learning (Gervasio *et al.* 2005). Over time, we envision the user being able to delegate more autonomy to the agent as it learns.

Example

Recall the example of conference travel and laptop purchase. Let c_1 be the candidate goal "Purchase a laptop", c_2 be the candidate goal "Attend AAAI", and c_3 the candidate goal "Attend AAMAS". Available funds are the resource r. Suppose c_1 and c_2 have led to the following adopted goals and intentions: g_1 with intention i_1: "Purchase a high-specification laptop using general funds"; and g_2 with intention i_2: "Attend AAAI and its workshops, staying in conference hotel". The agent estimates it is 90% through completion of intention i_1 and 25% through intention i_2; to change i_2 would incur a financial cancellation penalty.

Suppose the user tasks the agent with the new candidate goal c_3 to attend AAMAS in Europe. Thus, the current cognitive state is $S = \{B, \{c_1, c_2, c_3\}, \{g_1, g_2\}, \{i_1, i_2\}\}$. As before, suppose adopting c_3 is infeasible due to resource contention, i.e., insufficient general funds. The agent has several alternatives, including one or more of:

1. Do not adopt c_3 (i.e., don't attend AAMAS)
2. Drop c_1 or c_2 (i.e., laptop purchase or AAAI attendance)
3. Modify g_2 to attend only the main AAAI conference
4. Adopt a new candidate goal c_4 to apply for a departmental travel grant

Suppose the user has stated goal advice forbidding the agent from dropping any intention, thus disallowing the second alternative above. Finally, suppose the user also gave a high priority when tasking the agent with c_3.

According to its nature, the agent builds the optimization problem from the criteria it considers. In this case, suppose the agent's simple behavior is to neglect measures of commitment to existing intentions (i.e., here, the estimated percentage complete) and the cost of changing an intention, and to consider only the weight the user has given to candidate goals together with her goal advice. Despite the penalty incurred, the agent therefore seeks to move to an S' that corresponds to taking both the third and fourth options above: to modify g_2 and adopt g_4, i.e., $S' =$

$\{B, \{c_1, c_2, c_3, c_4\}, \{g_1, g_2, g_4\}, \{i_1, i_2\}\}$. Thus the agent is able to adopt g_3 also and maintain goal feasibility,

Notice how the combined cognitive process in the proposed framework enables the agent to have a global view of the interplay between goals and intentions.

Discussion and Future Work

Motivated by the need for cognitive models of agency suitable for agents that interact with humans, this paper introduced a BDI model of delegative assistance. The model distinguishes desires, candidate goals, and goals, and incorporates user guidance in the form of advice to influence the agent's cognitive processes. The model includes a combined mechanism for goal adoption and commitment revision, which is cast as a multi-criteria optimization problem constrained by the user's advice. We are in the process of operationalizing the delegative model within the SPARK agent framework, for use within CALO.

Delegative agents constitute one type of assistive agent. An important area for future work is to expand our model to encompass a broader set of human–agent interactions. One interesting direction is to support proactive identification of candidate goals by the agent on behalf of its user. So, for example, the agent may determine that it would be helpful to its user to initiate the collection of information from members of a project team to support the user in writing an upcoming annual report. The challenge here is for the agent to have sufficient understanding of the user's current desires and activities to act beneficially on her behalf. A second direction is to consider a more collaborative model of problem solving, in which the user and the assistant work together to perform tasks. This direction would involve merging our model of agency with prior ideas on collective problem solving (Cohen & Levesque 1991).

Last, we have largely left aside multiagent considerations in this paper. As noted, since CALO's user most likely works in a collaborative environment, CALO's rationality must include negotiation and cooperation with other agents. Advice can be seen as one means of encapsulating norms, not only of the human–agent, user–CALO interaction, but also the norms of multi-CALO interaction.

Acknowledgments. The authors thank D. Israel and D. Morley for helpful discussions. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA), through the Department of the Interior, NBC, Acquisition Services Division, under Contract No. NBCHD030010.

References

Allen, J.; Blaylock, N.; and Ferguson, G. 2002. A problem solving model for collaborative agents. In *Proc. of AAMAS'03*.

Ancona, D., and Mascardi, V. 2003. BCDI: Extending the BDI model with collaborations. In *Proc. of AAMAS'03 Workshop on Declarative Agent Languages and Technologies*, 146–161.

Choi, D.; Kaufman, M.; Langley, P.; Nejati, N.; and Shapiro, D. 2004. An architecture for persistent reactive behavior. In *Proc. of AAMAS'04*, 988–995.

Cohen, P. R., and Levesque, H. J. 1991. Teamwork. *Nous, Special Issue on Cognitive Science and Artificial Intelligence* 25(4).

Dastani, M., and van der Torre, L. 2002. Specifying the merging of desires into goals in the context of beliefs. In *Proc. of the First Eurasian Conf. on Advances in Information and Communication Technology (EurAsia ICT 2002)*, LNCS 2510, 824–831.

Dastani, M., and van der Torre, L. 2004. Programming BOID-plan agents: Deliberating about conflicts among defeasible mental attitudes and plans. In *Proc. of AAMAS'04*, 706–713.

Dignum, F.; Morley, D.; Sonnenberg, E.; and Cavedon, L. 2000. Towards socially sophisticated BDI agents. In *Proc. of ICMAS-2000*, 111–118.

Dignum, F.; Kinny, D.; and Sonenberg, E. 2002. From desires, obligations and norms to goals. *Cognitive Science Quarterly* 2(3–4):407–430.

Doyle, J., and Wellman, M. P. 1991. Impediments to universal preference-based default theories. *Artificial Intelligence* 49(1–3):97–128.

Georgeff, M. P., and Ingrand, F. F. 1989. Decision-making in an embedded reasoning system. In *Proc. of IJCAI'89*, 972–978.

Gervasio, M. T.; Moffitt, M. D.; Pollack, M. E.; Taylor, J. M.; and Uribe, T. E. 2005. Active preference learning for personalized calendar scheduling assistance. In *Proc. of the 2005 Intl. Conf. on Intelligent User Interfaces (IUI'05)*.

Levi, I. 1997. Advising rational agents. *Economics and Politics* 9(3):221–224.

Maheswaran, R.; Tambe, M.; Varakantham, P.; and Myers, K. 2003. Adjustable autonomy challenges in personal assistant agents: A position paper. In *Proc. of the AAMAS'03 Workshop on Agents and Computational Autonomy*, 187–194.

Morley, D., and Myers, K. 2004. The SPARK agent framework. In *Proc. of AAMAS'04*, 714–721.

Myers, K. L., and Morley, D. N. 2002. Conflict management for agent guidance. In *Proc. of AAMAS'03*, 1041–1042.

Myers, K. L., and Morley, D. N. 2003. Policy-based agent directability. volume 7 of *Multiagent Systems, Artificial Societies, and Simulated Organizations*. Berlin: Springer. chapter 9.

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

Pfeffer, A. 2005. Functional specification of probabilistic process models. In *Proc. of AAAI-05*, 663–669.

Rao, A. S., and Georgeff, M. P. 1991. Modeling agents within a BDI-architecture. In *Proc. of KR'91*, 473–484.

Rao, A. S., and Georgeff, M. P. 1995. BDI agents: from theory to practice. In *Proc. of ICMAS-95*, 312–319.

Rao, A. S., and Georgeff, M. P. 1998. Decision procedures for BDI logics. *J. Logic and Computation* 8(3):293–343.

Schut, M.; Wooldridge, M.; and Parsons, S. 2004. The theory and practice of intention reconsideration. *J. Experimental and Theoretical Artificial Intelligence* 16(4):261–293.

SRI International. 2005. CALO: Cognitive Assistant that Learns and Organizes. `www.ai.sri.com/project/CALO`.

Steuer, R. 1986. *Mulitple Criteria Optimization: Theory, Computation, and Application*. New York: Wiley.

Thangarajah, J.; Padgham, L.; and Harland, J. 2002. Representing and reasoning for goals in BDI agents. In *Proc. of the Australasian Conference on Computer Science*.

Winikoff, M.; Padgham, L.; Harland, J.; and Thangarajah, J. 2002. Declarative and procedural goals in intelligent agent systems. In *Proc. of KR'02*, 470–481.