Incorporating Planning and Reasoning into a Self-Motivated, Communicative Agent

Daphne Liu and **Lenhart Schubert**
Department of Computer Science
University of Rochester
Rochester, NY USA

Abstract

Most work on self-motivated agents has focused on acquiring utility-optimizing mappings from states to actions. But such mappings do not allow for explicit, reasoned anticipation and planned achievement of future states and rewards, based on symbolic knowledge about the environment and about the consequences of the agent's own behavior. In essence, such agents can only behave reflexively, rather than reflectively. Conversely, planning and reasoning have been viewed within AI as geared towards satisfaction of explicitly specified user goals, without consideration of the long-range utility of the planner/reasoner's choices. We take a step here towards endowing a self-motivated, utility-optimizing agent with reasoning and planning abilities, and show that such an agent benefits both from its knowledge about itself and its environment, and from exploiting opportunities as it goes. Our simulated simple agent can cope with unanticipated environmental events and can communicate with the user or with other agents.

Introduction

There is rather broad agreement in AI that general human-like intelligent behavior, apart from lower-level activities like perception and reflexes, is guided by planning and reasoning. However, planning and reasoning have traditionally been understood as aimed at the fulfillment of specified user goals, rather than as internally motivated by consideration of the long-range utility of the planner/reasoner's choices. Conversely, research on self-motivated agents has focused almost exclusively on acquiring utility-optimizing mappings from states to actions, without reasoned anticipation and planned attainment of future states and rewards based on symbolic knowledge about the environment and consequences of the agent's own behavior. These observations have motivated our work on *explicit self-awareness* (Sch05) and more thoughtful self-motivation in agents.

Here we present a self-aware and self-motivated agent that thinks ahead, plans and reasons deliberately, and acts reflectively, both by drawing on knowledge about itself and its environment and by seizing opportunities to optimize its utility. As an initial step towards a full conversation agent, our simple agent can communicate with the user or with other agents in addition to coping with unforeseen environmental events while acting opportunistically.

In the following sections, we discuss the notions of *explicit self-awareness* (Sch05) and thoughtful self-motivation, as well as how they are realized in our system. Then we outline our system and describe how our agent benefits from self-awareness, thoughtful self-motivation, and opportunistic choices. We conclude with a summary and a discussion of future work.

Explicit Self-Awareness

Explicit self-awareness was characterized by Schubert (Sch05) as being both human-like and explicit. Specifically, it is human-like in that an explicitly self-aware agent must have a well-elaborated human-like model of the world, including a model of itself and its relationships to the world. The self-model encompasses its beliefs, desires, intentions, knowledge, abilities, autobiography, the current situation, etc.; in addition, the agent must be capable of goal- and utility-directed reasoning and planning.

In addition to prescribing the aforementioned human-like capabilities, explicit self-awareness is explicit in three respects. First, an agent's representation of self-knowledge must be amenable to self-observation and use by the agent (and for engineering reasons, browsable and comprehensible to the designer). Second, explicit self-awareness must be conveyable by the agent, through language or other modalities. Third, the agent's self-knowledge must be amenable to inferences in conjunction with world knowledge.

Schubert (Sch05) enumerated reasons motivating the need for explicit self-awareness. First, given its bootstrapping potential with respect to meta-control, error recovery, autoepistemic reasoning, and learning of skills or facts, explicit self-awareness would help expand the frontiers of AI. Moreover, an explicitly self-aware agent can interact with human users in a transparent, natural, and engaging manner by having a shared context. Lastly, operational, explicitly self-aware agents can serve as exemplars of entities whose internal basis for self-awareness can be analyzed by consciousness theorists wishing to better understand self-awareness.

For additional discussions of explicit self-awareness, see (MS05; MS07; MS08) and (Liu08). The last elaborates on the contrast with other (weaker) notions of self-awareness as exhibited by self-monitoring agents (e.g., the metacognitive loop of Anderson and Perlis (AP05)), self-explaining agents (e.g., SHRDLU by Winograd (Win71)), global workspace systems (e.g., the opportunistic planning model by Hayes-Roth and Hayes-Roth (HRHR79)), and adaptive and robust goal-directed systems (e.g., an antibody system combatting viral intrusions). These conceptions of self-awareness either do not assume a self-model, or do not assume integration the self-model into general reasoning mechanisms.

Our conception of self-awareness has much in common with that of McCarthy, who proposes a formalization in terms of a mental situation calculus (McC95). He postulates that a machine must declaratively represent its mental states in order to introspect – observe and reason about its mental states, including beliefs, desires, intentions, knowledge, abilities, and consciousness. Schubert's proposal (Sch05) further specifies the knowledge representation and reasoning requirements for explicit self-awareness. In addition to a basic logical framework, these requirements include logical formulations of events, situations, attitudes, autoepistemic inferences, generic knowledge, and various metasyntactic devices such as axiom schemas, knowledge categorization, knowing or deriving a value, and experience summarization.

Our agent, dubbed ME for Motivated Explorer, to some extent meets the requirements for explicit self-awareness. ME knows specific facts about itself and the current situation, expressed as ground predications, as well as possessing general knowledge in the form of Horn-like clauses (but also allowing for reified propositions and questions). In effect, ME has a self-model that relates it to its simulated world and potentially the user. ME's knowledge of the current situation is initialized with its initial location, its possessions, geographical knowledge about the world, the current state facts about itself, and its propositional attitudes. For example, facts ($book\ book5$), ($owns\ ME\ book5$), ($knows\ ME\ (that\ (likes\ Grunt\ book5))$), and ($knows\ ME\ (whether\ (readable\ book5))$) in ME's knowledge base specify that ME owns the book $book5$, knows whether $book5$ is readable, and knows that entity $Grunt$ likes $book5$.

ME operates according to a plan of actions to be carried out, which ME dynamically modifies, evaluates for expected cumulative utility, and partially executes. Specifically, ME thinks ahead into the future and chooses to execute a seemingly best action. Such an action is one that constitutes the first action in some plan (with a limited horizon) that is judged to be executable from the current state, and in addition is anticipated to yield the highest cumulative utility among all such plans. Both the contemplated actions and the states they are expected to lead to can contribute positively or negatively to the anticipated utility of a plan. As actions are performed, ME's knowledge base, as well as the world state, will evolve accordingly. For instance, when ME obtains a book, the effect will be that ME has the book; this fact will enter ME's knowledge base, and as a result it will *know* that it has the book.

ME is introspective in three key respects. First, ME has knowledge of what operators it can invoke and what goals it can readily achieve in a given state. Second, ME can introspect about what it knows and doesn't know, and can handle propositional attitudes. Third, when ME performs an action, this is recorded in ME's history list of all actions and exogenous events that have occurred thus far in the world. ME's history list and knowledge base are both open to introspection, enabling ME to engage in more interesting question-answering with the human user and with other agents.

In a limited way, ME also meets the reasoning and communication requirements for explicitly self-aware agents. Apart from its ability to plan, it can also reason. In any given state, it performs bounded forward inference based on all of its current factual knowledge and all of its general quantified knowledge. One current limitation is that ME is excessively "skeptical", in the sense that it presumes to be false any ground predication that it cannot establish. Conversely, ME depends on a "commonsense inertia" assumption that whatever was true in the past and has not observably become false remains true.

Thoughtful Self-Motivation

A dynamic world poses three main types of challenges to a planning agent; namely, unexpected changes can arise in the form of unexpected action failures, unexpected threats, and unexpected serendipitous opportunities. In essence, these unexpected eventualities are due to the agent's incomplete or partial knowledge of the world. Thus the agent is necessarily confronted with the qualification and ramifications problems, that is, it simply does not know all the conditions for an action to succeed or all the possible consequences of an action, and so may experience unexpected outcomes. This is aggravated by the possibility of unpredictable exogenous events such as rain and fire. In the face of such indeterminacy, it is important that the agent act opportunistically in order to recover from (unexpected) failures, avoid (unexpected) threats, and pursue (unexpected) favorable opportunities in an appropriate and timely manner.

ME's opportunistic behavior is the byproduct of its constant, step-by-step striving towards maximum cumulative utility. For instance, suppose ME chooses to walk from home to school as that seems to be a best action to take. While walking from home to school, it may encounter a fire that makes the road unnavigable, at which point ME, undeterred, will do another lookahead into the future to select a best next action to take. Or, while walking from home to school, ME may see an unclaimed ten-dollar bill along the way, at which point it may pocket it if it finds doing so sufficiently pleasing (i.e., if doing so turns out to be the first step of the (cumulatively) most promising course of action).

ME is self-motivated in the sense that it has its own metrics of rewards (and penalties), and is driven by the "desire" to maximize cumulative rewards, rather than by some particular symbolic goal assigned to it, and to be pursued at all costs. Nonetheless, it pursues goals, to the extent that those goals promise high returns. ME's self-motivation is thoughtful, because of its grounding in reasoned look-ahead and evaluation. (Details concerning the look-ahead scheme are given in the following section.) Such deliberate self-motivation differs importantly from the impulsive self-motivation inherent in behavioral robots and reinforcement-learning agents, as most commonly understood. Broadly, such an agent functions in accord with a (preprogrammed or learned) *policy* that maps states of the world to the actions the agent should take in those states in order to maximize some overall reward criterion. However, neither the policy, nor the search for it, is guided by reasoning about future actions and situations, but instead both depend on the current state alone, and any past experience associated with it. In fact, reasoning typically is not an option, because states are typically not represented by symbolic descriptions, or in any case are not amenable to application of general planning and inference methods. The notions of beliefs, desires, and intentions are realized in only very elementary ways in such agents, if at all.

ME's self-motivation does not necessarily imply selfishness. While it may find certain self-involved states and actions rewarding (e.g., eating) or displeasing (e.g., being tired), and this will definitely affect its behavior, ME can also experience vicarious satisfaction, for example by answering the human user's questions, or, say, helping another agent in its world in some way. Notably, the anticipatory satisfaction in the look-ahead can also be used to implement curiosity, by making the acquisition of new knowledge, or going to as yet unvisited places, intrinsically satisfying for ME.

System Implementation

The knowledge, planning and behavior (both physical and dialog) of ME are programmed in LISP in a simulated world consisting of locations, connecting roads, and both animate entities (agents) and inanimate entities (objects) positioned at various locations. Some objects in the world may be consumable or portable and potentially useful to ME, while others might be harmful or mere obstacles. ME navigates the world interacting with the human user (in dialog) and with other entities in the world. All agents except ME are stationary and might be asked questions by ME and may supply things or information that ME wants upon request.

Creation of a world is enabled through commands for creating a road network, defining object types with various properties, and for placing instances of object types, with additional properties, at various locations in the network. Miscellaneous general knowledge can also be added. The additional properties of an instantiated entity include its associated objects (such as possessions or parts), and initial state facts about it, such as location or edibility, and for agents, propositional attitudes (beliefs, wants). ME's initial knowledge base contains the geographical knowledge about the world, general quantified conditional facts (with a conjunctive antecedent and a positive predication as consequent), and ME keeps a history list of all actions and exogenous events that have occurred so far in the simulated world. Examples of general knowledge might be properties of certain types of entities (e.g., that a sasquatch is an animate agent), or that certain conditions imply others (e.g., being asleep implies not being awake).

ME does not in general know the current facts, possessions, or propositional attitudes associated with other entities. However, all *non-occluded, local* facts about an entity become known to ME when ME is at the location of the entity. Occluded facts are determined by certain predicates (such as *hungry, knows*, or *contains*) being marked as occluded; as mentioned above, a fact with an occluded predicate is known initially only to the subject of the predication, if that subject is animate. For example, *hungry* might be an occluded predicate, but the subject $\langle term \rangle$ of a fact ($hungry\ \langle term \rangle$) is assumed to know that it is hungry whenever this is true. Moreover, ME may discover occluded knowledge via appropriate actions. For instance, the action *open* applied to a box followed by *read-message* may cause ME to know the contents of the message if the box has a message in it.

ME's action types have a list of parameters, a set of preconditions, a set of effects, and an associated value. One of the more unusual features is that both preconditions and effects allow for procedural evaluation or simplification, once all parameters are bound. In this way quantitative preconditions and effects can be handled quite effectively, as can side-effects such as ME producing a printed answer. As an example, consider the following operator *sleep* with formal fatigue and hunger level parameters $?f$ and $?h$, respectively:

```
(setq sleep
   (make-op :name 'sleep :pars '(?f ?h)
   :preconds '((is_at ME home)
               (is_tired_to_degree ME ?t)
               (>= ?f 0.5)
               (is_hungry_to_degree ME ?h)
               (> ?f ?h)
               (not (there_is_a_fire)))
   :effects '((is_tired_to_degree ME 0)
              (not (is_tired_to_degree
                     ME ?f))
              (is_hungry_to_degree ME
                     (+ ?h 2)))
   :time-required '(* 4 ?f)
   :value '(* 2 ?f)))
```

From ME's perspective, if it is at home, is more tired than hungry, is at least of fatigue level 0.5, and there is no fire, then it can sleep for a duration given by $(*\ 4\ ?f)$

and, as a result, it will relieve its fatigue at the expense of increasing its hunger level by 2. Performing an instantiated *sleep* action will afford ME a net increase of $(* \; 2 \; ?f)$ in its cumulative utility.

Instantiating an operator requires replacing its formal parameters with actual values through unifying the preconditions with facts in ME's current knowledge base, and such an instantiated action is considered applicable in the current state. At all times, ME maintains a plan comprised of a sequence of instantiated actions. Planning is accomplished by forward search from a given state, followed by propagating backward the anticipated rewards and costs of the various actions and states reached, to obtain a seemingly best sequence of actions. The forward search is constrained by a search beam, which specifies the allowable number of branches and the allowable operators for each search depth. Informed by this projective forward search, ME will then execute the first action of the seemingly best plan, and update its knowledge accordingly (in effect, observing non-occluded facts, including ones that have become false, in its local environment).

The simulated world is a dynamic one in which exogenous events such as fire and rain can spontaneously begin and end with some probability at each time step; therefore, unexpected changes can arise in the form of unexpected action failures, unexpected threats, and unexpected serendipitous opportunities. For example, a fire may start and disrupt ME's travel, or ME may scratch a lottery coupon and find that it has won one million dollars. Since the world is only partially known and partially predictable from ME's perspective, and since actions (such as traveling) can take multiple time steps, with the possibility of interference by exogenous events, we need to model "actual" actions in ME's world separately from ME's (STRIPS-like) conception of those actions.

The following is the stepwise version *sleep.actual* of the *sleep* operator:

```
(setq sleep.actual
   (make-op.actual :name 'sleep.actual
                   :pars '(?f ?h)
   :startconds '((is_at ME home)
                 (is_tired_to_degree ME ?t)
                 (>= ?f 0.5)
                 (is_hungry_to_degree ME ?h)
                 (> ?f ?h))
   :stopconds '((there_is_a_fire)
                (is_tired_to_degree ME 0))
   :deletes '((is_tired_to_degree ME ?#1)
              (is_hungry_to_degree ME ?#2))
   :adds '((is_tired_to_degree ME
             (- ?f (* 0.5
                     (elapsed_time?))))
           (is_hungry_to_degree ME
             (+ ?h (* 0.5
                     (elapsed_time?)))))))
```

The start conditions as given by *startconds* are the same except for removal of the (*there_is_a_fire*) formula. Notably, the actual action will continue for another time step if and only if neither of its stop conditions as given by *stopconds* is true in the current state. If at least one of them is true in the current state, then the action will immediately terminate. Otherwise, the current state and ME's knowledge base will be updated with ME's lower fatigue level and higher hunger level.

ME currently has various operators at its disposal, enabling it to answer the user's yes/no questions and wh- questions, to walk, sleep, eat, drink, ask other agents whether something is true, play, read, withdraw money from a bank, buy something from a store, and work and save money. Moreover, via operator (*listen!*), the user can signal to ME that a question or assertion (in symbolic form, not in English at this point) is about to be sent, and ME will "hear" and save that assertion or question, and potentially respond. Since answering questions has been assigned a high utility, ME will prefer to answer questions and verbalize its responses (as English sentences printed on the screen). Alternatively, for diagnostic reasons, the user is also provided with the ability to peer into ME's knowledge base and obtain the answer to a question immediately, without having to ask ME a question.

Preliminary Results

To empirically demonstrate the benefits of explicit self-awareness and of opportunistic (but still thoughtful) behavior in a self-motivated agent, we have created scenarios allowing some initial ablation tests. Here we describe some as yet incomplete attempts to investigate ME's performance (in terms of cumulative utility) with and without self-knowledge, and with and without opportunistic tendencies.

In all scenarios, there are four locations *home*, *grove*1, *plaza*1, and *company*1, with road *path*1 of length 2 connecting *home* and *grove*1, *path*2 of length 3 connecting *home* and *plaza*1, and *path*3 of length 2 connecting *grove*1 and *company*1 in the simulated world. Agent ME's knowledge base is initialized to reflect that it is at *home*, is not tired, has a thirst level of 4, has a hunger level of 2, and knows that *applejuice*1 is potable and at home. Object *pizza*1 is edible and at *plaza*1. Object *applejuice*1 is potable and at *home*. Agent *guru* knows whether *applejuice*1 is potable and whether *pizza*1 is edible.

In addition, ME has a variety of operators at its disposal, enabling it to answer the user's yes/no questions and wh-questions, walk, sleep, eat, drink, ask other agents whether something is true, play, read, buy something from a store, and work and save money. Also there are two types of exogenous events, namely fire and rain. Provided there is no rain, a spontaneous fire has a 5% chance of starting; once it has started, it has a 50% chance of stopping, and it also goes out as soon as there is rain. Spontaneous rain has a 33% chance of starting; once it has started, it has a 25% chance of stopping.

In normal operation, ME will gain rewards from the actions it performs (e.g., roaming, eating, or answering user queries) and the states it reaches (e.g., not being hungry, thirsty, or tired). One rather trivial way to ablate self-awareness is to eliminate all first-person knowledge such as (*is_at ME home*), (*is_hungry_to_degree ME* 2), etc., without altering operator definitions. In such a case, ME can no longer confirm the preconditions of its own actions, since these all involve facts about ME; thus, it is immobilized. But a more meaningful test of the effect of ablating first-person knowledge should replace ME's conceptions of its operators with ones that make no mention of the agent executing them, yet are still executable, perhaps with no effect or adverse effects in actuality, when actual preconditions unknown to ME are neglected. We would then expect relatively haphazard, unrewarding behavior, but this remains to be implemented.

A more interesting test of the advantages of self-awareness would be one focused on first-person *metaknowledge*, e.g., knowledge of type (*knows ME* (*whether* (*edible pizza*1))). Intuitively, given that ME can entertain such knowledge, and there are ways of finding out whether, for instance, (*edible pizza*1), ME should be able to plan and act more successfully than if its self-knowledge were entirely at the object- (non-meta-) level. In fact, this is why our scenarios include the above kind of meta-precondition in the definition of the *eat* operator, and why they include a guru who can advise ME on object edibility. Our experimentation so far successfully shows that ME does indeed ask the guru about the edibility of available items, and is thereby is enabled to eat, and hence to thrive. However, ablating meta-level self-knowledge, much as in the case of object-level self-knowledge, should be done by replacing ME's conceptions of its operators with ones that do not involve the ablated predications (in this case, meta-level predications), while still keeping the actual workings of operators such as *eat* more or less unchanged. So in this case, we would want to make an *eat*-simulation either unrewarding or negatively rewarding if the object to be eaten is inedible. This would surely degrade ME's performance, but this also remains to be confirmed.

To investigate the effects that ablation of opportunistic behavior has on ME's cumulative utility, we will focus our attention on one representative scenario. Initially, ME is at home feeling hungry and thirsty, knows *applejuice*1 at home to be potable, but does not know any item to be edible. To find out about the (only) edible item *pizza*1, ME must walk to *grove*1 and ask *guru*. With sufficient lookahead, having knowledge about *pizza*1 will incline ME to walk to *plaza*1 and eat *pizza*1 there. To entirely suppress ME's opportunistic behavior, we designate eating *pizza*1 as ME's sole goal and make ME uninterested in any action other than asking *guru* to acquire food knowledge, traveling to reach *guru* and *pizza*1, and eating *pizza*1.

In this case, none of ME's actions are disrupted by any spontaneous fire, and upon accomplishing its goal of eating *pizza*1 after 18 steps, ME achieves a cumulative utility of 66.5. The sequence of actions and events, each annotated with its time of occurrence in reverse chronological order, is as follows:

```
((EAT PIZZA1 PLAZA1) 17), (FIRE 15), ((WALK HOME PLAZA1
PATH2) 14), ((WALK HOME PLAZA1 PATH2) 12), ((WALK GROVE1
HOME PATH1) 9), (RAIN 9), (FIRE 8), ((WALK GROVE1 HOME
PATH1) 5), (RAIN 5), ((ASK+WHETHER GURU (EDIBLE PIZZA1)
GROVE1) 3), (FIRE 2), ((WALK HOME GROVE1 PATH1) 1),
((WALK HOME GROVE1 PATH1) 0), (RAIN 0).
```

With its opportunistic behavior restored, ME thinks ahead into the future and chooses to execute a seemingly best action at each step, achieving a higher cumulative utility of 80.5 after 18 steps. The higher cumulative utility comes from ME's better choices of actions (e.g., drinking when thirsty); specifically, it is a direct result of ME's seizing the initial opportunity to drink the potable *applejuice*1 to relieve its thirst, and ME can see and exploit such an opportunity because it is not blindly pursuing any one goal but rather is acting opportunistically. This case is shown below; no spontaneous fire disrupts any of ME's actions, and ME also finds out about *pizza*1 and eventually eats it.

```
((EAT PIZZA1 PLAZA1) 17), (RAIN 16), ((WALK HOME PLAZA1
PATH2) 15), ((WALK HOME PLAZA1 PATH2) 13), (RAIN 13),
((WALK GROVE1 HOME PATH1) 11), (RAIN 11), ((WALK GROVE1
HOME PATH1) 10), (RAIN 9), ((ASK+WHETHER GURU
(EDIBLE PIZZA1) GROVE1) 8), (FIRE 7), ((WALK HOME GROVE1
PATH1) 0) 6), ((WALK HOME GROVE1 PATH1) 5), (RAIN 5),
(FIRE 2), ((DRINK 4 APPLEJUICE1 HOME) 0), (RAIN 0).
```

These preliminary results indicate that ME can indeed benefit from both self-awareness and opportunistic, thoughtful self-motivation. While the results are encouraging, we are planning on doing systematic evaluations of our hypotheses, as we will outline in the concluding section.

Conclusion

We have presented an explicitly self-aware and self-motivated agent that thinks ahead, plans and reasons deliberately, and acts reflectively, both by drawing on knowledge about itself and its environment and by seizing opportunities to optimize its cumulative utility.

We have pointed out that deliberate self-motivation differs significantly from the impulsive self-motivation in behavioral robots and reinforcement-learning agents, driven by policies acquired through extensive experience (or through imitation), but not guided by symbolic reasoning about current and potential future circumstances. Our approach can be viewed as an integration of two sorts of agent paradigms – behavioral (purely opportunistic) agents on the one hand, and planning-based (goal-directed) agents on the other. Agents in the former paradigm focus on the present state, using it to choose an action that conforms with a policy reflecting past reward/punishment experience. Agents in the latter paradigm are utterly future-oriented, aiming for some goal state while being impervious to the current state, except to the extent that the current state supports or fails to support steps toward that future state. Some work in cognitive robotics (e.g., (TWN04; FFL04)) and in autonomous agents in computer games (e.g., (DEVG08)) intersects our approach, in that moves are chosen on the basis of the expected value of a sequence of moves (for instance, for a player in a Robocup world, or a person walking in a crowd, avoiding collisions). But generally these agents either are focused on externally supplied goals, or use feature-based rather than logical representations of states, and so cannot truly reason about them.

Additionally, we noted that our agent to some extent meets the knowledge representation and reasoning requirements (Sch05) for explicitly self-aware agents. Its ability to handle propositional attitudes (and in that sense metaknowledge) are particularly relevant to that point. Its self-knowledge, world knowledge, and introspection enable it to create and evaluate possible plans; furthermore, ME uses its current factual knowledge in any given state to perform bounded forward inference.

There are issues to address in our future work. First, ME is excessively skeptical in its presumption that ground predications are false whenever they cannot easily be established. Ignorance should not equal negation of knowledge. For example, not knowing if there is food does not mean there is no food; instead, being hungry and not knowing if there is food should prompt the agent to find out if there is food, that is, the agent should still consider pursuing eating. Some of this can probably be handled alternatively with "($\neg$) know-whether" propositions in ME's knowledge base. If ME knows whether ϕ but ϕ is not inferrable by or known to ME, then ME can conclude $\neg\phi$. If ϕ is not known to ME and ME does not know whether ϕ, then this might motivate ME to find out whether ϕ holds.

Eventually, degrees of uncertainty should be allowed for in ME's knowledge. If ME has definite negative knowledge of a precondition, then ME certainly should not consider pursuing the action. On the other hand, if it is still possible that a precondition currently not known to be satisfiable might be true, and if ME would like to pursue this action, then ME should aim to prove or disprove this precondition. To make ME less skeptical, we can specify that if ME has been to a location, then any non-occluded facts at that location that are not known by ME to be true are false; otherwise, no such assumption should be made.

Ultimately, we envisage an explicitly self-aware and self-motivated conversation agent with knowledge- and suggestion-driven dialogue behavior. The agent's behavior is ultimately driven by a planning executive that continually augments, modifies and partially executes a "life plan" that guides all of the agent's deliberate actions, whether physical, verbal or mental. Given the large number of possible dialogue moves corresponding to particular dialogue states, it is desirable to guide such a continually evolving planning executive by "suggestions" (certain kinds of if-then rules) triggered by the current situation. Such suggestions could be extremely helpful in inclining the agent to particular actions in particular situations.

By way of systematic demonstration, we would ideally want to show that our agent comes closer to reaping the maximum attainable cumulative utility than does either a purely opportunistic one or a solely goal-directed one. This raises the challenge of computing what the maximum attainable cumulative utility actually is in a given scenario. A possible approach to computing this maximum may be exhaustive forward search, as far into the future as possible. We would also want to explore probabilistic versions of such evaluations, in not-entirely-predictable worlds.

References

M. L. Anderson and D. R. Perlis. Logic, self-awareness and self-improvement: The metacognitive loop and the problem of brittleness. *Journal of Logic and Computation*, 15(1):21–40, 2005.

J. Dinerstein, P.K. Egbert, D. Ventura, and M. Goodrich. Demonstration-based behavior programming for embodied virtual agents. *Computational Intelligence*, 24(4):235–256, 2008.

E. Ferrein, C. Fritz, and G. Lakemeyer. On-line decision-theoretic golog for unpredictable domains. In *Proceedings of the 27th German Conference on Artificial Intelligence (KI2004)*, pages 322–336, Ulm, Germany, 2004.

B. Hayes-Roth and F. Hayes-Roth. A cognitive model of planning. *Cognitive Science*, 3:275–310, 1979.

D. H. Liu. A survey of planning in intelligent agents: from externally motivated to internally motivated systems. Technical Report TR-2008-936, Department of Computer Science, University of Rochester, June 2008.

J. McCarthy. Making robots conscious of their mental states. In *Machine Intelligence 15*, pages 3–17, 1995.

F. Morbini and L.K. Schubert. Conscious agents. Technical Report TR-2005-879, Department of Computer Science, University of Rochester, September 2005.

F. Morbini and L. Schubert. Towards realistic autocognitive inference. In *Logical Formalizations of Commonsense Reasoning, Papers from the AAAI Spring Symposium*, pages 114–118, Menlo Park, CA, March 26-28 2007. AAAI Press.

F. Morbini and L. Schubert. Metareasoning as an integral part of commonsense and autocognitive reasoning. In *AAAI-08 Workshop on Metareasoning*, pages 155–162, Chicago, IL, July 13-14 2008. Revised version in M. T. Cox & A. Raja (eds.), *Metareasoning: Thinking about Thinking*, MIT Press, to appear.

L. Schubert. Some knowledge representation and reasoning requirements for self-awareness. In M. Anderson and T. Oates, editors, *Metacognition in Computation: Papers from the 2005 AAAI Spring Symposium*, pages 106–113. AAAI Press, 2005.

M. Tacke, T. Weigel, and B. Nebel. Decision-theoretic planning for playing table soccer. In *Proceedings of the 27th German Conference on Artificial Intelligence (KI2004)*, pages 213–225, Ulm, Germany, 2004.

T. Winograd. Procedures as a representation for data in a computer program for understanding natural language. Technical Report 235, Massachusetts Institute of Technology, February 1971.