

MODELING TIME AND META-REASONING IN DIALOGUE VIA ACTIVE LOGIC

D. Perlis, K. Purang, D. Purushothaman, C. Andersen, D. Traum

University of Maryland

Computer Science Dept.

College Park, MD 20782

perlis,kpurang,darsana,cfa,traum@cs.umd.edu

Abstract

We describe our work on human-computer cooperative dialogue, that uses special “active” logics in an attempt to more closely – and more usefully – model human behavior. This effort pays attention to a number of issues, including Gricean pragmatics, logical non-omniscience, reasoning with inconsistent data, limited memory, blending in new information during reasoning, and models of human communication competence. Time is a chief unifying feature in our approach to these issues.

1 Introduction

In the paper[17] we outlined a general approach to automated dialogue, based on a variety of considerations including ideas from cognitive psychology. The resulting thesis there was that meta-dialogue can play a fundamental role in making effective dialogue possible, and that a proper treatment of time in turn plays a fundamental role in making effective meta-dialogue possible. Our work since then has aimed at testing the promise of that approach. (We have defined conversational adequacy as the ability to carry on free-ranging conversation, even on unfamiliar topics — in which case the dialogue may take on the form of one speaker learning by asking definitional and clarificational questions.)

More specifically, we proposed that the brittleness of current dialogue systems stems in large measure from their inability to “rethink” the immediate past course of the dialogue, note errors, and incorporate that information smoothly into the constantly evolving dialogue. In particular, the ability to reason explicitly about the actual words in the dialogue, when they were uttered and by whom, and how they were meant, appears crucial to initiating and carrying out repairs of miscommunication. This includes reasoning about spelling (for instance when two words sound much alike) or when one agent has misspelled a word; and about definitions (“how are you using that word?”); about more general conveyed meanings (“what did you intend in asking that?”); about obligations as well as the physical world; and about resource limitations (memory and attention). This is a tall list, and our aim has not been to achieve it all, but rather to explore the particular gains that may arise from the extensive use of time-based meta-dialogue.

In this paper we report on that work (including a somewhat broader context of our work on commonsense reasoning), and then present our aims for future work.

In our work on dialogue over the past four years, we have applied (and suitably modified) active logics to deal with such issues as Gricean pragmatics, logical non-omniscience, reasoning with inconsistent data, limited memory, blending in new information during reasoning, and models of human communication competence. This paper, half project-report and half position-paper, describes some of this work and why we think it is a promising approach.

Humans are real agents in a real world, subject to the constraints that reality imposes. Among these are constraints of time and space. Our work in natural-language human-computer dialogue attempts to model these real-world constraints, and to address traditional concerns of pragmatics in dialogue within that same model.

Many of these concerns involve metareasoning; thus we use a logic-based language so that inferences can be performed. The logics we use are “active logics”, developed in order to provide an evolving notion of time that can be reasoned about, even as it changes. This is important for many applications, including dialogue where sensitivity to the evolving dialogue itself (its immediate past history, who said what when, etc) is critical to understanding, responding, and making on-line corrections as part of the same dialogue.

We present a number of aspects of such reasoning, give examples of them occurring in dialogue, and suggest how active logic might provide computational tools for their successful performance in an automated system. Our current effort is centered on a Maryland version of the Rochester TRAINS-96 system [8] (which is a human-computer collaborative route-planning system that involves communication by dialogue (typed or spoken). We have restructured the dialogue manager around a meta-reasoning inference engine called ALMA (“Active Logic MACHine”). Our aim here is to test the hypothesis that meta-reasoning can play a major role in making human-computer dialogue more fluid and effective. In this we have been motivated by work in various aspects of cognitive psychology, including memory [1], deductive reasoning [3, 4, 5, 24, 15], everyday reasoning [2, 9, 18, 20], meta-reasoning [13], communication [22, 19] and relevance [7, 21].

All of these stress human abilities that, in AI terms, involve reasoning with a meta-theory or similar KRR tool. We have *not* focused our work on adhering closely to any one cognitive model or behavior, or even class of models or behaviors; but we have given, and continue to give, serious attention to theoretical and empirical work in cognitive psychology, in order to better assess what constraints the real world really does put on full-bodied agents.

We begin with a description of active logic.

2 Active Logic

Active logics were developed as a means of combining the best of two worlds – inference and reactivity – without giving up much of either. This requires a special evolving-during-inference model of time. The motivations for this were twofold: all physically realizable agents must deal with resource limitations the world presents, including time limitations; and people in particular have limited memories [1] and processing speeds, so that inference goes step by step rather than instantaneously as in many theoretical models of rationality. A consequence of such a resource-limited approach is that agents are not omniscient: there is no moment at which an agent has acquired all logical consequences of its beliefs. This not only must be so for real agents (and hence for humans) but it also is an advantage when the agent has contradictory beliefs (as real agents will often have, if only because their information sources are too complex to check for consistency). In this case, an omniscient and logically complete reasoner is by definition swamped with all sentences of its language as beliefs, with no way to distinguish safe subsets to work with. By contrast, active logics, like human reasoners, have at any time only a finite belief set, and can reason about their present and past beliefs, forming new beliefs as they do so; and this occurs even when their beliefs may be inconsistent. (See [11] for details.)

2.1 Deadline-coupled Reasoning

One key application of active logic is to deadline-coupled reasoning. An approaching deadline must be factored into one’s reasoning, even seen as an evolving part of that reasoning, rather than as a separate concern outside the reasoning process. Thus the remaining time (deadline – current-time) shrinks steadily as one attempts to find a solution to the problem. Yet, most automated planners

treat the planning as if it were instantaneous or as if the deadline does not come closer during planning; it was a major aim of active logics to overcome this limitation and to model the “to-do, doing, done” nature of real plans and actions, within an inference system that itself can serve as the rational portion of such an agent. That is, plans evolve from being on a to-do list, to be underway, to being completed (as plans), and then again as plans-in-execution they proceed from to-do, doing, done; and it behooves an agent to have such state information available so that as events unfold, and new information (such as messages from another agent) arrives, these can be factored into the ongoing reasoning appropriately¹.

Deadlines also occur in dialogue, though often less dramatically. A speaker who pauses generally expects a response before long. A delayed response of fifteen seconds is quite noticeable, puzzling, even irritating, suggesting that the listener wasn’t listening, or is being insulting. Thus the listener will usually recognize having a social obligation to respond quickly, and her deliberation as to what to say should not prevent actually saying something longer than a few seconds, even if the immediate response is “Hmm, let me think a bit more about this.” We are building into our system the KR tools, within active logics, that will allow us to code such behaviors in terms of time-sensitive metareasoning.

2.2 Formalism

The formal changes to a first order logic required here are, in some respects, quite modest. The language can be that of a first-order logic, perhaps augmented with names for expressions to facilitate meta-reasoning.²

The principal change is that inference rules become time-sensitive. The most obvious case is that of reasoning about time itself, as in the rule

i:	Now(i)

i+1:	Now(i+1)

The above indicates that from the belief (at time i) that the current time is in fact i , one concludes that it *now* is the later time $i + 1$. That is, time does not stand still as one reasons.

Technically, an active logic consists of a first-order language, a set of time-sensitive inference rules, and an observation-function that specifies an environment in which the logic “runs”. Thus an active logic is not pure formalism but is a hybrid of formal system and embedded inference engine, where the formal behavior is tied to the environment via the observations and the internal monitoring of time-passage (see [6] for a detailed description).

2.3 An example

For instance, suppose you are driving en route to the airport and planning details of your route as you go. You wonder whether to take the more direct but more heavily traveled road, or another. There are many facts to consider (time of day, day of week, radio traffic and weather reports) and

¹Planning systems such as PRS [12] deal with similar issues, but the approach is procedural. We believe greater flexibility can be obtained from a logical approach such as ours.

²Time-sensitive meta-reasoning is the key to deadline-planning, contradiction-repair, and most of the features distinguishing active logics from other logics. (This meshes well with work in human metacognition; see for instance [13].) Collaborative systems must reason not only about their own beliefs and actions but also about the beliefs and actions of their partners. This is a yet more involved kind of meta-reasoning, studied both in traditional formalisms and in active logics.

many implications to ferret out (the radio is not broadcasting any traffic news, but it may be due to lack of such news or to their obsession with announcing a baseball game, etc). You quickly realize that your flight will be gone before you can figure out ramifications to all these subtleties. So you decide to stop worrying about the best of all possible routes, and instead content yourself with any one that seems likely to work.

Using time sensitive active-logic inference rules similar to the rule for *now* (section 2.2), deadline-coupled reasoning has been formalized and applied to planning problems (see [14]) where time of plan-enactment is crucial.

In this example the reactivity is not to external events but rather to the universal event of time-passage vis-a-vis one's own reasoning. One can conceptualize this externally in terms of looking at a clock but this is not necessary or particularly helpful. On the other hand, external events are often quite important, as we discuss later.

3 Consequences of Active Logics

Psychological modeling is behind most of the research goals of active logics. These are not intended to specifically model precise details of human behavior (not to model precise human memory characteristics, for instance) but rather to obtain useful ideas from human behavior and incorporate them into our system design. Thus the short-term/long-term memory distinction, ability to change topic, to participate in conversation about topics one knows little about (and learn as a result) are all design considerations that motivate our work. Our view is that, since humans are the only known intelligent agents in existence, we should avail ourselves as much as we can about human behavior in the attempt to design such systems.

Time, change of mind, recall, and attention to past events, all play large roles in human intelligence and conversation; our aim is to see to what extent these and other notions can be incorporated into a formal framework where the semantics is relatively clear. We present some of the consequences of active logics in that respect.

3.1 Time Situatedness

Active logics are able to react to incoming information (including dialogue utterances by a collaborative partner) while reasoning is ongoing, blending new inputs into its inferences without having to start up a new theorem-proving effort. Thus, any helpful communications of a partner (or user) – whether as new initiatives, or in response to system requests – can be fully integrated with the system's evolving reasoning. Similarly, external observations of actions or events can be made during the reasoning process and also factored into that process.

To return to the airport example: you are driving en route to the airport and planning details of your route as you go. Then your car gets a flat tire. Rather than complete your original plans, it is time to make major revisions, in fact even to change one's goals for the time being.

Gricean pragmatics is another area where active logics can be brought to bear [16]. For instance, during dialogue, a listener's beliefs can change markedly and even reverse themselves. This happens in the case of presuppositional inference, as in the dialogue

A: Are the roses in the fridge? [presupposition: there are roses and
there is a fridge]

B: There are no roses. [above presupposition for roses has

to be withdrawn]

Similarly, implicatures can be reversed, as in:

A: Are the roses fresh?

B1: They are in the fridge... [they are probably fresh]

B2: ...but they are not fresh [above implicature withdrawn]

In both cases, an inference made at one point is denied later on. This change must be computed by the participants in real time, since it affects the continuation of the same dialogue. Thus the reasoning, the uttering, and the metareasoning about the dialogue, all go on in time that itself is kept updated as part of that metareasoning. Active logics have been applied to this successfully, where traditional logics are not able to represent the changes and also carry out the reasoning involved.

3.2 Self-reference

In active logic the flat tire above can be represented in terms of observations; and the reasoning simply goes on with this new information. There is no executive subsystem that turns off the route planner midstream and starts up a new planning action. Rather there is a single stream of reasoning which can monitor itself by looking backwards at one moment to see what it has been doing in the past, including the very recent past. If the previous few steps in some way conflict with new information, then the next few steps can be devoted to sorting out enough of the apparent mismatch to allow a decision as to how to proceed. All of this is carried out in the same inferential process as the original planning, without the need for level upon level of meta-reasoners. This is not to say that there is no metareasoning here, but rather that it is “in-line” metareasoning, all at one level. The advantages of this are (i) simplicity of design, (ii) no infinite regress, and (iii) no reasoning time at higher levels unaccounted for at lower levels.

Self-referential dialogue is very common, as in “This conversation is getting nowhere,” or “Can you hear me?” As Grice has pointed out, in some sense almost every utterance is self-referential, in that part of its intended meaning is that its meaning be conveyed by that same utterance. This appears to set up an expectation as to how listeners will react, and which can be modeled over time in active logic: the utterance, its being heard, the expectation of listener reaction, the actual reaction, revised expectations, and so on.

A potential disadvantage is the possibility of vicious self-reference. This matter is a topic of current investigation.

3.3 Inconsistency Tolerance

Another major advantage of such time-sensitive in-line metareasoning is that inconsistency in one’s beliefs need not cause serious problems in general. The reason is largely that given above: a conflict in the reasoner’s beliefs can be noted by the reasoner as a new belief, and the latter can lead to a decision to encapsulate the conflicting beliefs so they do no harm. Now this cannot be a fully general process, since identifying contradictions is at best semi-decidable. However, deeply hidden contradictions may usually do little harm; and so we have concentrated on inference rules for “direct” contradictions, that is, belief pairs that surface in the form P and $\neg P$. Miller [11]

discusses this in more detail and proves a theorem providing a rather general case in which such in-line metareasoning can cope with direct contradictions.

Again, in dialogue, clashes of information are common; one person contradicts another (or herself), and listeners must deal with this, deciding which (if either) contradictand to tentatively accept. This occurred in the examples of presuppositional and implicatural inferences (section 3.1).

3.4 Representing Changing State

Another feature that comes directly out of the time-coupled nature of active logics is their ability to represent the evolving status of reasoning and actions. For instance, it is easy to write an inference rule that updates at each time step whether a particular plan is currently being started, is already underway, or is completed. More detail than this, such as how long the plan execution has been going on, is also readily inferred. This is important for various purposes, such as:

- avoiding re-initiation of a plan already underway
- assessing whether one is spending too much time on that goal
- deciding whether adopting another plan might be better

3.5 Words and Meanings

The meta-representation in active logics allows us to represent words separately from their meanings. We can therefore also reason about them. This is important in language processing when we want to clarify the meaning of a word, for instance. [11] discusses this in the context of the problem of misidentification.

4 Problems

Although an active logic does not model an omniscient reasoner, the problem of resource limitations is still very important. Additional special inference rules will be needed for limited memory and attention. Some preliminary work on this has been done, but not enough to implement in a dialogue system. The belief set in an active logic is always finite, but can become quite large, which is not in keeping with known bounds on human working memory; nor is it reasonable from a computational perspective. This raises the issue of relevance: if memory is limited, which items are most important to retain? And that in turn involves issues such as topic and intention. Active logics appear to be adequate mechanisms for introducing such further constraints; the hard part is deciding precisely what form those constraints should take.

5 Current Work

Our work to date has brought us to being able to process several simple types of requests, like

```
‘‘Send the Boston train to New York’’  
‘‘Where are the trains?’’  
‘‘Where is Metroliner?’’
```

Often, of course, these requests are misprocessed, and that is where our main research interests lie: in devising methods to recognize and repair such errors.

We are currently designing the knowledge representation and inference tools that will allow us to process complex dialogue interactions of the “do X ... no, do Y” sort. A particularly interesting case occurs when X and Y are the same e.g.:

```
‘‘Send the Boston train to New York...  
...no, send the Boston train to New York’’
```

Here the above statements are by a human user to a train-delivery planning system; the system has misunderstood “Boston train” in the initial part of the utterance, and instead moved another train (perhaps the “Brockton” train, if this is a case of misunderstood speech; or perhaps the train originally at Boston rather than the one there now, if this is a case of poorly disambiguated reference).

In order for the system to properly interpret the correction in the latter portion of the above utterance, it must come to recognize that “no” is not a change of mind on the user’s part (as it might have been), nor is it an incoherent self-contradiction by the user (don’t send it and do send it), but rather an implicit correction of the intervening action taken by the system (sending, say, the Brockton train instead of the Boston train).

Such misunderstandings, and the need for their repair, are very common in human dialogue, and even more so in human-computer dialogue. We have already sketched a solution for various forms of miscommunication in the domain of the TRAINS-96 dialogue system, following the general pattern “do X ... no, do Y”, as reported in [23]. In that paper five distinct types of miscommunication errors in “do X ... no, do Y” utterances (including the one above) were isolated, and a uniform set of tools for their solution method was proposed.

6 Future Work

6.1 The Immediate Future

The next step is to implement and test our prototype solution for the various “do X ... no, do Y” forms examined in (Traum and Andersen, 1999). This will allow us to analyze the implications of such methods towards the more general problem of miscommunication repair.

This step will involve: (i) implementation of the specific KR and inference patterns sketched in (Traum and Andersen, 1999); (ii) implementation of an efficient interface between those tools and our existing inference engine; (iii) adding the spoken language front-end to the dialogue system and (iv) evaluating the system’s resulting behavior compared to the pre-existing TRAINS-96 system and user’s expectations.

Evaluation will be done in two ways:

- by comparing system output in each case to the intended “correct” output (task success).

Thus, in the ‘‘do X ... no, do X’’ example, the system should, proceed to send the train currently at Boston, to New York.

- by analyzing the internal generality of the reasoning.

Thus in the ‘‘do X ... no, do X’’ sample discourse above we would determine whether the system

- (a) discovered the apparent incoherence (do X, don't do X, do X)
- (b) examined various potential interpretations
- (c) resolved these into the one determined to have the greatest
 plausibility (which would be different for the original
 instruction 'do X' and the correction 'no, do X')

On the basis of the above, we will be better able to formulate our approach to longer-range efforts in the processing of miscommunication errors. The kinds of errors that can occur in "do X ... no, do Y" are fairly representative ones. But these cases are made a bit easier by the corrective help given in the second ("no...") part; that is why we are starting out with such cases.

Sometimes the context allows for largely automatic error-processing; the "do X ... no, do Y" cases tend to be of this sort, since the requester already indicates much of the solution in the second part of the utterance. However, even then there can be ambiguities, as we noted above, and this is where automated meta-reasoning can come into play.

Thus after testing various types of such cases in the above study, we can then proceed to other types of errors (e.g., miscommunications due to spelling errors, errors of attention, and definitional miscommunications) both in the "do X ... no, do Y" cases and also in less user-guided settings where the system must decide whether an error has been committed, whether to initiate a one-sided repair or seek assistance, and so on.

Another short-term goal is the introduction of speech input. When there is spoken input, rather than typed, new difficulties arise, and in particular there are new types of ambiguities and corresponding miscommunications to be repaired. To resolve some of these ambiguities, (especially when context reasoning does not help to remove the ambiguity), it is imperative for the system to seek assistance from the user in some cases whereas in others this may not be needed; eg:

USER: Send the Boston train to Stanton.

SYS: [sends Boston train to Stanton]

USER: Send the new Stanton train to Maryland.

Now, the system would have to find out whether the user meant the train at New Stanton (assuming there is such a town), or the newly arrived train at Stanton. In typed input the use of capital letters or quote marks can allow the user to clarify this at the outset; but in speech this is conveyed through subtle tonal inflections and not easily captured by speech-recognition systems.

USER: Send the Boston train to New York first. Send the Maryland train to Boston.

Now, the system may have to find out from the user which action is to be performed first, since in speech the period after "first" is not available.

In addition to the above ambiguities, the agent should be highly reactive to external inputs to take care of 'panic inputs' from the user as illustrated by the next example.

USER: Have you sent the Boston train to Maryland?

SYS: [starts to send Boston train to Maryland]

USER: No! Don't do that!

Here the first two words were not caught by the speech-recognizer (a common occurrence), and as a result an inappropriate action is begun, and must be stopped as soon as possible.

Active Logic appears appropriate to deal with situations of the above sorts, because of its inherent reactivity and meta-reasoning capabilities.

6.2 The Long Range

One major ongoing application of active logics is that of building a "conversationally adequate" dialogue agent— one can view such an agent as having the ability to learn in McCarthy's sense of

advice-taking, via conversation. Preliminary work on applying active logics to problems in language processing has been done [10, 16], and we have proposed an abstract view of how we would build such a conversationally adequate agent [17]. We view metareasoning to be a crucial part of that type of agent and believe that active logics are well suited for that. The focus now is on precisely specifying and implementing another piece of a conversationally adequate agent.

The long-range research plan envisions the following steps:

1. process a few simple queries via voice input, requiring the use of negative introspection (that is, the reasoner coming to know that it does not know some fact), in the context of the Rochester TRAINS system rewritten to employ active logic.
2. handle apparent contradictions as in “Send the Boston train to NYC... [here the system does something]... No, send the Boston train to NYC” where the system has misunderstood the first request, as we described earlier.
3. understand and generate spelling in metadialogue (“The word I am saying is spelled ‘xyz’...”)
4. definitions and/or definitional queries (“What does the word ‘xyz’ mean?”)
5. more general metadialogue: You said, I said, you meant, I meant, etc
6. reasoning about the physical and social worlds
7. focus, attention and memory
8. conversational adequacy

References

- [1] A. D. Baddeley. *Human memory: theory and practice*. Allyn & Bacon, 1990.
- [2] M. Chapman. Everyday reasoning and the revision of belief. In *Mechanisms of everyday cognition*. 12th Biennial conference on lifespan developmental psychology, L. Erlbaum Associates, 1993.
- [3] P. W. Cheng and K. J. Holyoak. Pragmatic reasoning schemas. *Cognitive Psychology*, 17:391–416, 1985.
- [4] P. W. Cheng, K. J. Holyoak, R. E. Nisbett, and L. M. Oliver. Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18:293–328, 1986.
- [5] L. Cosmides. The logic of social exchange: Has natural selection shaped how humans reason? studies with the wason selection task. *Cognition*, 31:187–276, 1989.
- [6] J. Elgot-Drapkin and D. Perlis. Reasoning situated in time I: Basic concepts. *Journal of Experimental and Theoretical Artificial Intelligence*, 2(1):75–98, 1990.
- [7] R. Elio and F. J. Pelletier. On relevance in non-monotonic reasoning: some empirical studies. In *Relevance Working notes AAAI94 Fall Symposium*, 1994.
- [8] George M. Ferguson, James F. Allen, Brad W. Miller, and Eric K. Ringger. The design and implementation of the TRAINS-96 system: A prototype mixed-initiative planning assistant. TRAINS Technical Note 96-5, University of Rochester, 1996.

- [9] K. Galotti. Approaches to studying formal and everyday reasoning. *Psychological Bulletin*, 105:331–351, 1989.
- [10] J. Gurney, D. Perlis, and K. Purang. Interpreting presuppositions using active logic: From contexts to utterances. *Computational Intelligence*, 1997.
- [11] M. Miller. *A View of One’s Past and Other Aspects of Reasoned Change in Belief*. PhD thesis, Department of Computer Science, University of Maryland, College Park, Maryland, 1993.
- [12] K. L. Myers. *User Guide for the Procedural Reasoning System*. SRI International,, Menlo Park, CA, 1997.
- [13] T. O. Nelson. *Metacognition: Core readings*. Allyn & Bacon, 1992.
- [14] M. Nirkhe, S. Kraus, M. Miller, and D. Perlis. How to (plan to) meet a deadline between *now* and *then*. *Journal of logic computation*, 7(1):109–156, 1997.
- [15] M. Oaksford and N. Chater. Reasoning theories and bounded rationality. In K. I. Manktelow and D. E. Over, editors, *Rationality: psychological and philosophical perspectives*, pages 31–60. Routledge, 1993.
- [16] D. Perlis, J. Gurney, and K. Purang. Active logic applied to cancellation of gricean implicature. In *Working notes, AAAI 96 Spring Symposium on Computational Implicature*. AAAI, 1996.
- [17] D. Perlis, K. Purang, and C. Andersen. Conversational adequacy: Mistakes are the essence. *International Journal of Human Computer Studies*, pages 553–575, 1998.
- [18] J. M. Puckett, L. K. Pollina, J. S. Laipple, R. H. Tu nick, and F. H. Jurden. A survey of research in everyday cognition: ten views. In *Mechanisms of everyday cognition*. 12th Biennial conference on lifespan developmental psychology, L. Erlbaum Associates, 1993.
- [19] S. Savignon. *Communicative Competence: Theory and Classroom Practice*. Addison-Wesley, 1983.
- [20] J. D. Sinnott and J. Cook. An overview– if not a taxonomy– of “Everyday Problems” used in research. In J. D. Sinnott, editor, *Everyday problem solving: theory and applications*. Praeger, 1989.
- [21] Dan Sperber and Deirdre Wilson. *Relevance: communication and cognition*. Blackwell, Cambridge, Massachusetts, 1995.
- [22] Elaine Tarone. Some thoughts on the notion of communication strategy. *TESOL Quarterly*, 15(3):285–295, 1981.
- [23] D. Traum and C. Andersen. Representations of dialogue state for domain and task independent meta-dialogue, 1999. Submitted to IJCAI99 Workshop: Knowledge And Reasoning in Practical Dialogue Systems.
- [24] P. C. Wason and D. W. Green. Reasoning and mental representation. *Quarterly Journal of Experimental Psychology*, 36A:597–610, 1984.