ELABORATION TOLERANCE

John McCarthy
Computer Science Department
Stanford University
Stanford, CA 94305
jmc@cs.stanford.edu
http://www-formal.stanford.edu/jmc/

1997 Dec 14, 6:10 p.m.

Abstract

A formalism is *elaboration tolerant* to the extent that it is convenient to modify a set of facts expressed in the formalism to take into account new phenomena or changed circumstances. Representations of information in natural language have good elaboration tolerance when used with human background knowledge. Human-level AI will require representations with much more elaboration tolerance than those used by present AI programs, because human-level AI needs to be able to take new phenomena into account.

The simplest kind of elaboration is the addition of new formulas. Next comes changing the values of parameters. Adding new arguments to functions and predicates represents more of a change. However, elaborations not expressible as additions to the object language representation may be treatable as an addition at a meta-level expression of the facts.

Elaboration tolerance requires nonmonotonic reasoning, although nonmonotonic reasoning is not the focus of this article. Representing contexts as objects in a logical formalism that can express relations among contexts should also help.

We use the missionaries and cannibals problem and about 20 variants as our *Drosophila* in studying elaboration tolerance in logical AI.

The present version has only some parts of a situation calculus formalization. However, the English language elaborations listed are enough to serve as a challenge to logical AI formalisms claiming elaboration tolerance.

1 Introduction

In several papers, e.g. [McC88] and [McC89], I discussed the *common sense informatic situation* and contrasted it with the information situation within a formal scientific theory. In the latter, it is already decided what phenomena to take into account. In the former, any information possessed by the agent is available and potentially relevant to achieving its goals. *Elaboration tolerance* seems to be the key property of any formalism that can represent information in the common sense informatic situation.

Elaboration tolerance[1] is the ability to accept changes to a person's or a computer program's representation of facts about a subject without having to start all over. Often the addition of a few sentences describing the change suffices for humans and should suffice for computer programs.

Humans have considerable elaboration tolerance, and computers need it to reach human-level AI. In this article we study elaboration tolerance in terms of logical AI. However, researchers pursuing other AI methodologies will also have to face the problem of elaboration tolerance; maybe they just haven't noticed it yet. The relation to*belief revision* will be discussed briefly in section 8.

Humans represent information about the world in natural language and use background knowledge not ordinarily expressed in natural language. [2] The combination of linguistic and non-linguistic knowledge is what gives us humans our elaboration tolerance. Unfortunately, psychological research hasn't yet led to enough understanding of the background knowledge, so it is hard to study elaboration tolerance in humans. However, it is easy to give plenty of examples of human elaboration tolerance, e.g. those in this article.

[1]The concept was first mentioned in [McC88].

[2]The non-linguistic background knowledge has been emphasized in connection with physical skills by Hubert Dreyfus and others [Dre92], but there is important non-linguistic knowledge also when the skill is purely symbolic. Even though a mathematician or a stock broker operates in a purely symbolic domain, he still cannot verbalize his full set of skills.

The *Drosophila*[3] of this article is the missionaries and cannibals problem (MCP).

After describing the original MCP, we give a large number of elaborations. Humans tolerate these elaborations in the sense that we can use the sentences expressing one of the elaborations to get a modified problem. People will agree on what the modified problem is and will agree on whether a proposed solution is ok.

Then we consider logical formalizations of the original problem and discuss which elaborations different formalisms tolerate. Our goal—not achieved in this article—is a formalism for describing problems logically that is as elaboration tolerant as English and the associated background knowledge. However, some logical languages are more elaboration tolerant than others.

2 The Original Missionaries and Cannibals Problem

The missionaries and cannibals problem (abbreviated MCP):

> **Three missionaries and three cannibals come to a river and find a boat that holds two. If the cannibals ever outnumber the missionaries on either bank, the missionaries will be eaten.**
>
> **How shall they cross?**

We call this original version of the problem MCP0.

Saul Amarel proposed [Ama71]: Let a state (mcb) be given by the numbers of missionaries, cannibals and boats on the initial bank of the river. The initial situation is represented by 331 and the goal situation by 000.

Most AI texts that mention the problem accept this formulation and give us the solution:

$$331 \rightarrow 310 \rightarrow 321 \rightarrow 300 \rightarrow 311 \rightarrow 110 \rightarrow 221 \rightarrow 020 \rightarrow 031 \rightarrow 010 \rightarrow 021 \rightarrow 000.$$

[3] *Drosophilas* are the fruit flies that have been used by geneticists to study inheritance since 1910. Their short generation times, large chromosomes and the ability to keep 1,000 of them in a bottle make them valuable, even though the *Drosophilas* of today are no better than those of 1910. The utility of suitable *Drosophilas* for scientific research in AI needs to be emphasized, because of a recent fad for demanding that all research promise a practical payoff on a three year schedule. They aren't getting their payoffs and are learning much less than a more scientific approach would get them.

If this representation is used the state space has 32 elements some of which are forbidden and two others are unreachable. It is an elementary student exercise to write a program to search the space and get the above sequence of states, and people are always solving it without a computer or without even a pencil. Saul Amarel [Ama71] points out that this representation has fewer states than a representation with named missionaries and cannibals.

What more does this problem offer AI?

If one indeed begins with the Amarel representation, the problem is indeed trivial. However, suppose we want a program that begins, as people do, with a natural language presentation of the problem. It is still trivial if the program need only solve the missionaries and cannibals problem. The programmer can then cheat as much as he likes by making his program exactly suited to the MCP. The extreme of cheating is to make a program that merely prints

$$331 \to 310 \to 321 \to 300 \to 311 \to 110 \to 221 \to 020 \to 031 \to 010 \to 021 \to 000.$$

Readers will rightly complain that this cheats, but it isn't clear what does and doesn't count as cheating when a method for solving a single problem is asked for.

The way to disallow cheating is to demand a program that can solve any problem in a suitable set of problems. To illustrate this we consider a large set of elaborations of MCP. It won't be trivial to make a program that can solve all of them unless the human sets up each of them as a state space search analogous to the original MCP. We demand that the program use background common sense knowledge like that about rivers and boats that is used by a human solver.

We skip the part about going from an English statement of the problem to a logical statement for two reasons. First, we don't have anything new to say about parsing English. Second, we don't yet have the logical target language that the parsing program should aim at. Progress toward establishing this language is the goal of the paper.

The problem is then to make a program that will solve any of the problems using logically expressed background knowledge. The background knowledge should be described in a general way, not specifically oriented to MCP and related problems.

This much was already proposed in **Programs with Common Sense**[4] [McC59]. What is new in the present paper is spelling out the idea of *elaboration tolerance* that was distantly implicit in the 1959 paper. We require a formulation of MCP that readily tolerates elaborations of the problem and allows them to be described by sentences added to the statement of the problem rather than by surgery on the problem. English language formulations allow this, but the Amarel-type formulations do not. AI requires a logical language that allows elaboration tolerant formulations.

We begin with examples of English language elaboration tolerance.

3 Some Elaborations Expressed in Natural Language

1. The boat is a rowboat. (Or the boat is a motorboat). This elaboration by itself should not affect the reasoning. By default, a tool is usable. Later elaborations make use of specific properties of rowboats.

2. The missionaries and cannibals have hats, all different. These hats may be exchanged among the missionaries and cannibals. In all the elaborations mentioned below, exchanging hats is an action irrelevant to crossing the river. There are two demands on the reasoner. Epistemologically, whatever reasoning that establishes a plan for crossing the river without the hats should be valid with the hats. This includes any nonmonotonic reasoning.

   Heuristically, the problem may not be trivial. Why should it be obvious that exchanging hats is of no use? Certainly we can make elaborations in which it is of use, e.g. we can assert that if the smallest missionary wears the hat belonging to the largest missionary, the largest cannibal won't eat him even if they go together. We can ask for this, however. It should be possible to tell a problem solver: Look for a solution that has no hat change actions. After that, the reasoner should find the solution as easily as it would if hats were never mentioned.

3. There are four missionaries and four cannibals. The problem is now unsolvable. In ordinary logic, adding sentences that there are four of

[4] http://www-formal.stanford.edu/jmc/mcc59.html

each produces a contradiction. Belief revision systems ought to make the correct change. It seems to me that people take a metalinguistic stance, just saying "Change the numbers of missionaries and cannibals to four", thus regarding the original statement of the problem as an object. Actually what is regarded as an object is the *sense* of the original statement, since people ordinarily don't remember the words used.

Proofs of impossibility take the following form. Choose a predicate formula $\phi(s)$ on situations. Show $\phi(S0)$ and $(\forall s)(\phi(s) \rightarrow \neg Goal(s))$. Also show

$$(\forall s\ a)(\phi(s) \rightarrow Bad(Result(a,s)) \vee \phi(Result(a,s))).$$

Thus you can't get out of the situations satisfying ϕ, and the goal isn't included. The simplest $\phi(s)$ is a disjunction of specific locations of the missionaries and cannibals in the reachable situations, but this disjunction is long, and it is very likely possible to do better.

4. The boat can carry three.

5. There is an oar on each bank. One person can cross in the boat with just one oar, but two oars are needed if the boat is to carry two people. We can send a cannibal to get the oar and then we are reduced to the original problem. [5] The formalism also needs to admit the reasoning that getting the oar reduces the problem to MCP0. A formalism using preconditions can accept this elaboration as just adding a precondition and adding facts about the locations of the oars in $S0$.

6. The boat leaks and must be bailed concurrently with rowing. Elaboration tolerance requires that treating a concurrent action be a small change in the statement of the problem, and this will show the limitations of some versions of situation calculus.

[5]It was not mentioned before that the boat was a rowboat. Once oars are mentioned, it is a Gricean implicature that the boat is a rowboat. The philosopher Paul Grice [Gri89] studied what can be inferred from statements under the assumption that the person posing the problem is not trying to be misleading. That the boat is a rowboat follows, because the speaker should have said so if it wasn't.

7. The boat may suffer damage and have to be taken back to the left bank for repair. This may happen at any time. This requires that the formalism permit splitting the event of crossing the river into two parts.

8. There is a bridge. This makes it obvious to a person that any number can cross provided two people can cross at once. It should also be an *obvious* inductive argument in the sense of McAllester [McA91]. This is a straightforward elaboration in situation calculus formalisms, since adding the bridge is accomplished just by adding sentences. There is no need to get rid of the boat unless this is part of the elaboration wanted.

9. There is an island. Then any number can cross, but showing it requires inductive arguments. Though inductive, these arguments should be obvious. Defining the three stages—moving the cannibals to the island, moving the missionaries to the opposite bank and then moving the cannibals to the opposite bank—is an easy three step problem, provided moving the sets of missionaries and cannibals can be regarded as tasks. Whether the elaboration is easy depends on the original representation.

10. Only one missionary and one cannibal can row. The problem is still solvable. Before this elaboration, we did not need to distinguish among the missionaries or among the cannibals. An elaboration tolerant language must permit this as an addition.

11. The missionaries can't row. This makes the problem impossible, since any solution requires two missionaries in the boat at some time. The formalism must admit the statement and proof of this lemma.

12. If the biggest cannibal is isolated with the smallest missionary, the latter will be eaten. A solution to the basic problem can be specialized to avoid this contingency. We have the Gricean implicature that the cannibals aren't all the same size, and need to have language for referring to an individual as the biggest cannibal and not just language to refer to him by name.

13. The biggest cannibal cannot fit in the boat with another person. The problem is solvable. However, if the biggest missionary cannot fit in the boat with another person the problem becomes unsolvable. We

can imagine having to elaborated in the direction of saying what sets of people can fit in the boat.

14. One of the missionaries is Jesus Christ. Four can cross. Here we are using cultural literacy. However, a human will not have had to have read Mark 6:48–49 to have heard of Jesus walking on water.

15. There are four cannibals and four missionaries, but if the strongest of the missionaries rows fast enough, the cannibals won't have gotten so hungry that they will eat the missionaries. This could be made precise in various ways, but the information is usable even in vague form.[6]

16. There are four missionaries and four cannibals, but the cannibals are not hungry initially, and the missionaries have a limited amount of cannibal food. They can tell if a cannibal is hungrier than he was and can avoid trouble by giving the food to the cannibal who has got hungrier. This requires comparing a situation and a successor situation.

17. Three missionaries alone with a cannibal can convert him into a missionary. The problem for elaboration tolerance is to change a predicate that doesn't depend on situation or time to one that does. If we used a sorted logic with missionaries and cannibals as distinct sorts, the intolerance might be embedded in the logic itself.

18. The probability is 1/10 that a cannibal alone in a boat will steal it. We can ask what is the probability that a given plan will succeed, say the Amarel plan. The formalism of [McC79a] treats *propositions* as objects. Using that formalism $\mathbf{Pr}(p) = 1/10$ can be expressed for any proposition p. I see at least two problems. The language of propositions as objects needs to be rich enough to express notions like the probability of a cannibal stealing the boat. The second problem is that we need to be able to assert independence or joint distributions without letting the entire formalism be taken over by its probabilistic aspects. In MCP0, cannibals have to be alone in the boat several times. We can write a formula that states that probabilities are independent by default.

19. There are two sets of missionaries and cannibals too far apart along the river to interact. The two problem should be solvable separately

[6] "Pull, pull, my good boys", said Starbuck.—Moby Dick, XLVIII

without considering interleaving actions at the two sites. If the two problems are different elaborations, the work required and the length of the proof should be the sum of the lengths for the separate problems plus a small constant.

There are N sites along the river with identical conditions. The reasoning should be able to do one site, or a generalized site, and, with a constant amount of additional reasoning, say that all N crossings are the same.

4 Nonmonotonic reasoning

Elaboration tolerance clearly requires nonmonotonic reasoning. For example, elaborating MCP0 with a requirement for oars adds preconditions to the action of going somewhere in the boat. If oars are not mentioned, nonmonotonic reasoning prevents such additional preconditions.

However, it is still not clear how to formulate the nonmonotonic reasoning so as to obtain tolerance of a wide class of elaborations, such as those of section 3. We propose to use some variant of circumscription, but this still leaves open what is to be circumscribed.

[McC80] discusses several nonmonotonic aspects of the human understanding of MCP0. They all have a Gricean [Gri89] character. They all concern the non-existence of features of the problem that should have been mentioned, were they supposed to exist. What can be inferred from such contexts includes the *Gricean implicatures*. Very likely, the formal theory of contexts [McC93] can be used, but that is beyond the scope of this article.

Here are a few nonmonotonic inferences that some up.

- If there were a bridge, it should have been mentioned. A puzzle problem like MCP is given in a context.
- The river can't be forded, and there isn't an extra boat.
- There isn't a requirement for a permit or visa to cross the river.
- There is nothing wrong with the boat. In general, when a tool is mentioned, it is supposed to be usable in the normal way.
- The group of missionaries and cannibals is minimized. Mentioning that one of the missionaries is Jesus Christ will include him in the

number otherwise inferrable rather than adding him as an additional missionary. Of course, assertions about him in the general database should have no effect unless he is mentioned in the problem.

These kinds of nonmonotonic reasoning were anticipated in [McC80] and have been accomodated in situation calculus based nonmonotonic formalisms, although the Yale shooting problem and others have forced some of the axiomatizations into unintuitive forms.

The elaborations discussed in this article mainly require the same kinds of nonmonotonic reasoning.

5 A Typology of Elaborations

There are many kinds of elaborations a person can tolerate, and they pose different problems to different logical formalizations. Here are some kinds of elaborations.

adding preconditions, actions and objects The example of the oars adds a precondition to rowing and adds the action of picking up the oars. Several situation calculus and event calculus formalisms allow this—assuming sentences are added before the nonmonotonic reasoning is done.

irrelevant actors, actions and objects Sentences establishing the existence of such entities should not vitiate the reasoning leading to a solution.

changing a parameter This is needed when the numbers of missionaries and cannibals are changed from 3 to 4. In English, this is accomplished by an added sentence. Doing it that way in logic requires a suitable *belief revision* method as part of the basic logical formalism. At present we must use minor brain surgery to replace certain occurrences of the number 3.

specialization In one situation calculus formalization we have the action $Move(b1, b2)$. If there are guaranteed to be exactly two places, we can replace this action by $Move(b)$, regarding this as $Move(b, Opp(b))$, where $Opp(b)$ designates the opposite bank and satisfies $Opp(Opp(b)) = b$. We regard this kind of specialization as an easy kind of elaboration.

generalization Some of our elaborations can be composed of an generalization of the language—replacing a function by a function of more arguments, e.g. making whether a person is a cannibal or missionary situation dependent or replacing going from a bank b to the opposite bank $Opp(b)$ by going from $b1$ to $b2$. Many elaborations consist of a generalization followed by the addition of sentences, e.g. adding preconditions or effects to an action.

going into detail An event like the action of crossing the river is made up of subactions. However, the relation between an event and its subevents is not often like the relation between a program and its subroutines, because asserting that the action occurs does not imply specific subactions. Rowing is a detail of crossing a river when rowboats are used, but rowing is not a part of the general notion of crossing a river.. Bailing if necessary is another detail. Getting oars or a bailing can are associated details. There is more about this apparently controversial point in [McC95].

m...s and c...s as actors MCP and almost all of the elaborations we have considered take a god-like view of the actions, e.g. we send a cannibal to get an oar. We can also elaborate in the direction of supposing that the actions of cannibals and missionaries are sometimes determined by the situation. In this case, it may be convenient to use a predicate $Occurs(event, s)$ and let one possible event be $Does(C_1, Enter\text{-}Boat)$. The situation calculus treatment has to be altered and looks more like event calculus.

simple parallel actions If one of the missionaries is Jesus Christ, we can transport 4 missionaries and 4 cannibals. We get 3 cannibals on the far bank and one on the initial bank. Then two ordinary missionaries and Jesus cross, the ordinaries in the boat and Jesus walking on water. The rest of the solution is *essentially the same* as in MCP0. A missionary and a cannibal row back and now the remaining two missionaries cross. We then send a cannibal to ferry the remaining two cannibals. We haven't tackled the problem of being able to say "essentially the same" in logic.

The formalization must permit Jesus to cross in parallel with the other missionaries so that the missionaries are never outnumbered. This isn't the same as having Jesus cross as a separate action.

full parallelism This is what permits requiring that the boat be bailed.

events other than actions The simple $Result(a,s)$ doesn't allow for events other than actions. To handle them we have used [McC95] a predicate $Occurs(e,s)$ asserting that the event e occurs in the situation s. Then $result(e,s)$ can be used.

comparing different situations This works ok in situation calculus, but some other formalisms don't allow it or make it awkward. Thus we can have $s <_{better} result(e,s)$ to say that the situation is better after event e occurs. We may also want $result(a1,s) <_{better} result(a2,s)$, comparing the result of doing $a1$ with the result of doing $a2$.

6 Formalizing the Amarel Representation

We use logic and set theory to formalize MCP0 and call the formalization MCP0a. In this formalization we are not concerned with elaboration tolerance. My opinion is that set theory needs to be used freely in logical AI in order to get enough expressiveness. The designers of problem-solving programs will just have to face the fact.

$$States = Z4 \times Z4 \times Z2$$

$$\begin{aligned}&(\forall state)(Ok(state) \equiv \\ &\quad Ok1(P1(state), P2(state)) \wedge Ok1(3 - P1(state), 3 - P2(state)))\end{aligned}$$

Here $Z2 = \{0,1\}$ and $Z4 = \{0,1,2,3\}$ are standard set-theory names for the first two and the first four natural numbers respectively, and $P1$, $P2$ and $P3$ are the projections on the components of the cartesian product $Z4 \times Z4 \times Z2$.

Note that having used $3 - P1(state)$ for the number of missionaries on the other bank put information into posing the problem that is really part of solving it.

$$(\forall m\ c)(Ok1(m,c) \equiv m \in Z4 \wedge c \in Z4 \wedge (m = 0 \vee m \geq c))$$

$$Moves = \{(1,0),(2,0),(0,1),(0,2),(1,1)\}$$

$$
\begin{array}{ll}
\multicolumn{2}{c}{(\forall move\ state)} \\
(Result(move, state) = Mkstate(& P1(state) - (2P3(state) - 1)P1(move), \\
& P2(state) - (2P3(state) - 1)P2(move), \\
& 1 - P3(state)))
\end{array}
$$

$$(\forall s1\ s2)(Step(s1, s2) \equiv (\exists move)(s2 = Result(move, s1) \wedge Ok(s2)))$$

$$Attainable1 = Transitive\text{-}closure(Step)$$

$$Attainable(s) \equiv s = (3, 3, 1) \vee Attainable1((3, 3, 1), s)$$

Notice that all the above sentences are definitions, so there is no question of the existence of the required sets, functions, constants and relations. The existence of the transitive closure of a relation defined on a set is a theorem of set theory. No fact about the real world is assumed, i.e. nothing about rivers, people, boats or even about actions.

From these we can prove

$$attainable((0, 0, 0)).$$

The applicability of MCP0a to MCP must be done by postulating a correspondence between states of the missionaries and cannibals problem and states of MCP0a and then showing that actions in MCP have suitable corresponding actions in MCP0a. We will postpone this until we have a suitable elaboration tolerant formalization of MCP.

MCP0a has very little elaboration tolerance, in a large measure because of fixing the state space in the axioms. Situation calculus will be more elaboration tolerant, because the notation doesn't fix the set of all situations.

7 Situation Calculus Representations

At present I do not know how to write a situation calculus formalization that tolerates all (or even most) of the elaborations of section 3. Nevertheless, I think it is useful to give some formulas that accomplish some elaborations and discuss some issues of elaboration tolerance that these formulas present.

7.1 Simple situation calculus

We begin with some axioms in a formalism like that of [McC86] using a *result* function, a single abnormality predicate and aspects. This suffers from the Yale shooting problem if we simply minimize ab. However, as long as the problem requires only projection, i.e. predicting the future from the present without allowing premises about the future, *chronological minimization* of ab [Sho88] avoids the Yale shooting problem. It is certainly a limitation on elaboration tolerance to not allow premises about the future.

Here are some axioms and associated issues of elaboration tolerance.

The basic operation of moving some people from one bank to the other is conveniently described without distinguishing between missionaries and cannibals.

$$
\begin{aligned}
(1)\quad & \neg Ab(Aspect1(group, b1, b2, s)) \rightarrow \\
& \qquad Value(Inhabitants(b1), Result(Cross(group, b1, b2), s)) = \\
& \qquad\qquad Value(Inhabitants(b1), s) \setminus group \\
& \wedge \\
& \qquad Value(Inhabitants(b2), Result(Cross(group, b1, b2), s)) = \\
& \qquad\qquad Value(Inhabitants(b2), s) \cup group,
\end{aligned}
$$

where $\setminus$ denotes the difference of sets.

The fact that (1) can't be used to infer the result of moving a group if some member of the group is not at $b1$ is expressed by

$$(\exists x \in group)(\neg Holds(At(x, b1), s)) \rightarrow Ab(Aspect1(group, b1, b2, s)).$$

It's bad if the cannibals outnumber the missionaries.

$$
\begin{aligned}
(2)\quad & Holds(Bad(bank), s) \\
& \equiv \\
& 0 < Card(\{x | x \in Missionaries \wedge Holds(At(x, bank), s)\}) \\
& < Card(\{x | x \in Cannibals \wedge Holds(At(x, bank), s)\})
\end{aligned}
$$

and

$$(3)\quad Holds(Bad, s) \equiv (\exists bank) Holds(Bad(bank), s).$$

$Card(u)$ denotes the cardinality of the set u.

Some of the elaborations involve restrictions on who can row. All of them can make use of

(4) $\neg(\exists x)(x \in group \wedge Rower(x)) \rightarrow Ab(Aspect1(group, b1, b2, s)).$

The oar-on-each-bank elaboration can be expressed by conjoining to the above

$$\begin{array}{l} Card(group) > Card(\{x | Oar(x) \wedge Holds(In(x, Boat), s)\}) \\ \rightarrow Ab(Aspect1(group, b1, b2, s)), \end{array}$$

but this looks a bit *ad hoc.*

Many unique names axioms will be required. We won't list them in this version.

7.2 Not so simple situation calculus

The notion of Bad in the previous subsection avoids any actual notion of the missionaries being eaten. More generally, it avoids any notion that in certain situations, certain events other than actions will occur. We can put part of this back.

To say that the missionaries will be eaten if the cannibals outnumber them can be done with the formalism of [McC95].

$$\begin{array}{l} Holds(Bad(bank), s) \rightarrow \\ \quad (\forall x)(x \in Missionaries \\ \qquad \wedge Holds(At(x, bank), s) \rightarrow Occurs(Eaten(x), s)). \end{array} \tag{5}$$

As sketched in [McC95], the consequences of the occurrence of an event may be described by a predicate $Future(f, s)$, asserting that in some situation in the future of the situation s, the fluent f will hold. We can write this

(6) $Future(f, s) \rightarrow (\exists s')(s <_{time} s' \wedge Holds(f, s')),$

and treat the specific case by

(7) $occurs(Eaten(x), s) \rightarrow F(Dead\text{-}soon(x), s)$

To say that something will be true in the future of a situation is more general than using $Result$, because there is no commitment to a specific next situation as the result of the event. Indeed an event can have consequences at many different times in the future. The $Result(e, s)$ formalism is very convenient when applicable, and is compatible with the formalism of $Occurs$ and F. We have

$$(8) \quad \neg Ab(Aspect2(e, s)) \wedge Occurs(e, s) \rightarrow F((\lambda s')(s' = Result(e, s)), s),$$

where something has to be done to replace the lambda-expression $(\lambda s')(s' = Result(e, s))$ by a syntactically proper fluent expression.

Perhaps we can use the $result$ formalism after all. We suppose that $result(e, s)$ is a situation after all the events consequential to e have taken place. We then have one or more consequences of the form

$$(9) \quad \rightarrow Past(f, result(e, s)),$$

and these permit us to refer to the consequences of e that are distributed in time. I don't know how often it will pay to make use of this possibility.

8 Remarks and Acknowledgements

1. Some commonly adopted conventions in theories of actions interfere with elaboration tolerance. An example is identifying situations or events with intervals of time. You can get away with it sometimes, but eventually you will be sorry. For example, you may want to say that a good move is one that leads to a better situation with

$$good(a, s) \equiv s <_{good} result(a, s).$$

2. *Elaboration tolerance* and *belief revision* have much in common, but we are looking at the problem from the opposite direction from researchers in belief revision. Belief revision studies have mainly concerned the effect of adding or removing a given sentence, whereas our treatment of elaboration tolerance concerns what you must add or change to get the effect you want. Moreover, the effect of an elaboration can involve changing the first order language and not just replacing one expression in the language by another.

3. Elaboration tolerance is rather straightforward when the theory to be changed has the structure of a cartesian product, and the elaboration can be describes as giving some components of the product new values. [McC79b] discusses theories with cartesian product structures in connection with counterfactuals, and [McC62] discusses the semantics of assignment, i.e. the semantics of changing components of a state.

4. Murray Shanahan [Sha97] considers many issues of elaboration tolerance in his discussions of action formalisms. In particular, his solutions for the frame problem are considerably elaboration tolerant. I qualified the above, because I consider elaboration tolerance an open ended problem.

5. I suspect that elaboration tolerance requires a proper treatment of *hypothetical causality* and this involves *counterfactual conditional* sentences. Counterfactuals will be treated in a shortly forthcoming paper by Tom Costello and John McCarthy. For example, we need a non-trivial interpretation of "If another car had come over the hill while you were passing, there would have been a head-on collision" that is compatible with the fact that no car came. By non-trivial interpretation, I mean one that could have as a consequence that a person should change his driving habits, whereas no such conclusion can be reached from sentences of the form $p \rightarrow q$ when p is false.

6. We can distinguish between a formalism admitting a particular elaboration and the consequences of the elaboration being entirely determined. For example, the Jesus Christ elaboration could be given alternate interpretations and not just the one about his ability to walk on water.

This paper has benefitted from discussions with Eyal Amir, Tom Costello and Josephina Sierra. It is available, possibly improved, as http://www-formal.stanford.edu/jmc/elaboration.html.

References

[Ama71] Saul Amarel. On representation of problems of reasoning about action. In Donald Michie, editor, *Machine Intelligence 3*, pages 131–171. Edinburgh University Press, 1971.

[Dre92] Hubert Dreyfus. *What Computers still can't Do.* M.I.T. Press, 1992.

[Gri89] Paul Grice. *Studies in the Way of Words.* Harvard University Press, 1989.

[McA91] David McAllester. Some observations on cognitive judgements. In *AAAI-91*, pages 910–915. Morgan Kaufmann Publishers, July 1991. internet file ftp.ai.mit.edu:/pub/dam/aaai91a.ps.

[McC59] John McCarthy. Programs with Common Sense[7]. In *Mechanisation of Thought Processes, Proceedings of the Symposium of the National Physics Laboratory*, pages 77–84, London, U.K., 1959. Her Majesty's Stationery Office. Reprinted in McC90.

[McC62] John McCarthy. Towards a mathematical science of computation. In *Information Processing '62*, pages 21–28. North-Holland, 1962. Proceedings of 1962 IFIP Congress.

[McC79a] John McCarthy. First order theories of individual concepts and propositions. In Donald Michie, editor, *Machine Intelligence*, volume 9. Edinburgh University Press, Edinburgh, 1979. Reprinted in [McC90].

[McC79b] John McCarthy. Ascribing mental qualities to machines[8]. In Martin Ringle, editor, *Philosophical Perspectives in Artificial Intelligence.* Harvester Press, 1979. Reprinted in [McC90].

[McC80] John McCarthy. Circumscription—A Form of Non-Monotonic Reasoning[9]. *Artificial Intelligence*, 13:27–39, 1980. Reprinted in [McC90].

[McC86] John McCarthy. Applications of Circumscription to Formalizing Common Sense Knowledge[10]. *Artificial Intelligence*, 28:89–116, 1986. Reprinted in [McC90].

[7]http://www-formal.stanford.edu/jmc/mcc59.html
[8]http://www-formal.stanford.edu/jmc/ascribing.html
[9]http://www-formal.stanford.edu/jmc/circumscription.html
[10]http://www-formal.stanford.edu/jmc/applications.html

[McC88] John McCarthy. Mathematical logic in artificial intelligence. *Daedalus*, 117(1):297–311, 1988.

[McC89] John McCarthy. Artificial Intelligence, Logic and Formalizing Common Sense[11]. In Richmond Thomason, editor, *Philosophical Logic and Artificial Intelligence*. Klüver Academic, 1989.

[McC90] John McCarthy. *Formalization of common sense, papers by John McCarthy edited by V. Lifschitz*. Ablex, 1990.

[McC93] John McCarthy. Notes on Formalizing Context[12]. In *IJCAI-93*, 1993. Available on http://www-formal.stanford.edu/jmc/.

[McC95] John McCarthy. Situation Calculus with Concurrent Events and Narrative[13]. 1995. Contents subject to change. Reference will remain.

[Sha97] Murray Shanahan. *Solving the Frame Problem, a mathematical investigation of the common sense law of inertia*. M.I.T. Press, 1997.

[Sho88] Yoav Shoham. Chronological ignorance: Experiments in non-monotonic temporal reasoning. *Artificial Intelligence*, 36(3):279–331, 1988.

/@steam.stanford.edu:/u/ftp/jmc/elaboration.tex: begun Mon Sep 8 10:59:15 1997, latexed December 14, 1997 at 6:10 p.m.

[11]http://www-formal.stanford.edu/jmc/ailogic.html
[12]http://www-formal.stanford.edu/jmc/context.html
[13]http://www-formal.stanford.edu/jmc/narrative.html