CS 323, Fall 2000
Problem Set #3: Situation Calculus and the Frame Problem
Handout #10

Due in class on Tuesday, November 21, 2000. (Two weeks from today.)

The purpose of this problem set is twofold. One is to introduce you to situation calculus and solutions to the frame problem. The other is to give a description of a first order logic theorem prover called Otter, which we will use to write and prove logical theories. There are 9 problems.

Please write legibly! Imprecise, unclear, or illegible answers will not receive full credit.

The material in this handout was largely pilfered from [Sha97].

1 Situation Calculus

1.1 The Basics

Situation Calculus was introduced in [MH69] as an ontology of time used to represent actions and change. There are lots of flavors of situation calculus, as well as other ontologies of time. We will present two flavors of sitcalc in this section.

Situation Calculus has three sorts or types: situations, actions, and fluents.[1] A situation is "the complete state of the universe." [MH69] Actions connect situations, and fluents are like propositions. There is one predicate $Holds(f, s)$ which asserts that fluent f is true in situation s. So to say it is raining at situation $S0$, we write $Holds(Raining, S0)$.[2] $Result(a, s)$ denotes the situation resulting from performing the action a in situation s. Thus $Result(Toggle(Switch1), S0)$ is the situation resulting from toggling the switch $Switch1$ in situation $S0$.

The structure of situations can be thought of as a tree, with $Result(a, s)$ branching off new situations, one for each action a from a situation s, and $Holds(f, s)$ holding at each of the nodes. There are two ways of interpreting these situations further. One is where each situation is uniquely determined by its place in the tree of situations (the *arboreal* interpretation), and the other is where each situation is uniquely determined by what fluents are true in it (the *situation state* interpretation).

1.1.1 Arboreal Interpretation

In this notion, as mentioned above, each situation is defined by its place in the situation tree. To formalize this notion, we need the following two axioms:

$$\forall a_1, a_2, s_1, s_2. Result(a_1, s_1) = Result(a_2, s_2) \rightarrow (a_1 = a_2 \wedge s_1 = s_2) \quad (1)$$

[1] Technically, there is also a fourth sort of objects, which are passed as arguments to fluents. So for example A and B are objects while $On(A, B)$ is a fluent.

[2] In [MH69], fluents were defined as predicates that took a situation as an argument. Instead of $Holds(Raining, S0)$ we would say $Raining(S0)$. There are pluses and minuses to each representation: using $Holds$ allows one to quantify over fluents in first-order logic. (We can say things like $\forall f. Holds(f, s)$.) The minus is a philosophical objection, where the reification of propositions is denounced as "ad hoc."[Gal91] The process of changing the predicate $Raining(s)$ into an object $Raining$ which is passed as an argument to a predicate is known as *reification*.

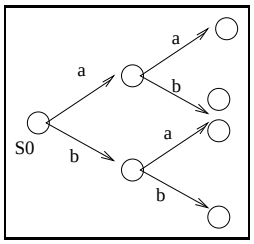

Figure 1: The Tree of Situations.

$$\forall a, s. \neg Result(a, s) = S0 \tag{2}$$

These two axioms are known as the *axioms of arboreality*. Each situation s is unique, and determined by the chain of actions leading from $S0$ to s. These are very similar to uniqueness of names axioms, or the Peano axioms. Equation 1 in the converse direction says that if either two situations or actions are different, then the situation resulting from each action is also different. Equation 2 says that there is no action and situation which can lead to $S0$.

Here is some more intuition about the formulas above. Equation 1 ensures that the graph of situations is really a tree – no branch is allowed to loop back to a situation upstream. (Try formulating a counter-example to see how this sentence prevents looping back.) Equation 2 prevents any situation from generating the special situation $S0$, or in other words, letting the entire tree loop back to its root. $S0$ is the special situation denoting the "first" situation. John calls this the "big bang" treatment of situations, as everything begins from $S0$.

The induction axiom is often mentioned:

$$\forall \phi. [(\phi(S0) \wedge \forall a, s. \phi(s) \rightarrow \phi(Result(a, s))) \rightarrow \forall s. \phi(s)] \tag{3}$$

The induction axiom says that all situations are generated by some sequence of actions from $S0$. (If we prove that any predicate Φ is true of $S0$, and is preserved from one situation to the next over any action a, then we've shown it true for every situation.) Another way to interpret it is that the set of situations is the smallest set S such that $S0 \in S$ and if $s \in S$ then so is $Result(a, s) \in S$.[3]

This axiom is used to prevent (even infinite) chains in our tree of situations, where the chain doesn't "stem" from $S0$ as depicted below:

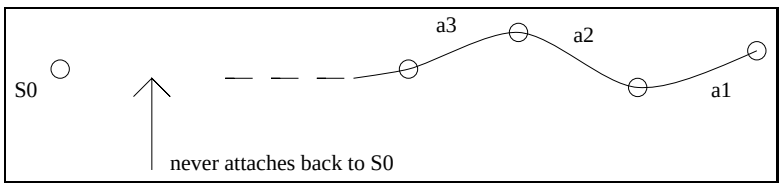

Figure 2: An Example of an Infinite Chain.

[3]We could rewrite Equation 3 as $\neg\exists\phi.(\phi(S0) \wedge \forall a, s. \phi(s) \rightarrow \phi(Result(a, s)) \wedge \exists s. \neg\phi(s))$. In other words, there is no property which, if true for $S0$ and preserved under $Result$, "leaves out" a situation.

1.1.2 Situation State

In the arboreal interpretation, situations are defined by the chain of actions which generate it from $S0$. An alternate interpretation is where a situation is determined by the fluents that are true in it. This is formalized by the following axiom:

$$\forall s_1, s_2 . s_1 = s_2 \iff [\forall f . (Holds(f, s_1) \iff Holds(f, s_2))] \tag{4}$$

Here we assert that if two situations have all the same fluents true (and false) in them, then they must be the same situation. In this case our tree of situations could become a graph. For example, if we had one action $Toggle$ which toggles a switch $Switch1$ on or off, then we would have the following depiction of situations:

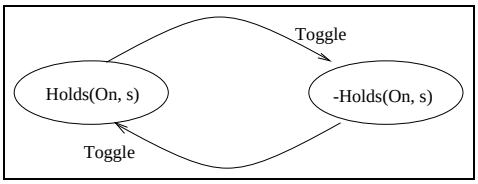

Figure 3: Situation States Collapse to a Cycle for Toggle.

Compare this to the arboreal treatment:

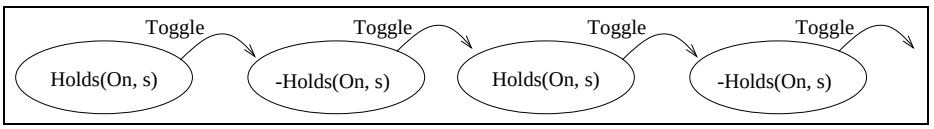

Figure 4: Arboreal Situations Do Not Collapse for Toggle.

Note that we would have add extra machinery to the situation-state version to account for time – we don't want time to be the same if we toggle a light switch on, off and then on again!

The general terminology is that states are objects which are uniquely determined by a set of fluents, while situations are the arboreal "branches" that are defined by a sequence of actions.[4]

Problem 1:[10 points] Prove that arboreal situation calculus defined by axioms 1 and 2 has no loops.

Problem 2:[10 points] Say we have A actions and F fluents in our language. How many unique situations are possible in the arboreal interpretation? What if we treated situations as states of fluents?

[4]Sometimes other connotations are given to states, in that they are determined by a fixed set of fluents, and thus can be represented as a vector of values. Situations are by contrast rich objects, whose properties can never be fully described. In this mind-set, situations are definitely more elaboration tolerant than states.

1.2 Limitations of Situation Calculus?

This discrete, tree-like representation for reasoning about actions and change seems clear and intuitive. Some skeptics point out it is also very limiting. Objections are that sitcalc cannot represent concurrent actions, continuous change, or actions where the order is unknown. It can't talk about divisible actions, or delayed effects of actions. It turns out that with a little extra machinery, each of these can be represented, as shown in [GLR91].

Problem 3:[10 points] Pick one of the solutions in [GLR91] and explain it in your own words. This paper is available in class or outside Gates 218.

1.3 The Frame Problem

The *frame problem* arises when formalizing change in a domain. To describe change, one has to write an axiom for each action a, describing under what preconditions it succeeds, along with what is changed. A sample axiom is below:

$$\forall x, y, s. Holds(Clear(x), s) \wedge Holds(Clear(y), s) \wedge x \neq y \wedge x \neq Table \rightarrow \\ Holds(On(x, y), Result(Move(x, y), s)) \tag{5}$$

The antecedent of the implication contains the preconditions for the action $Move(x, y)$ to succeed, and in the resulting situation $Result(Move(x, y), s)$, we know that x is on top of y. Formulas like this are known as *effect axioms*, for obvious reasons.

It seems easy – to formalize a domain, just take each action, decide what its preconditions and effects are, and then just write down a set of effect axioms. The problem is, the fact that the other fluents not affected by the action do not change also has to be formalized. For example, moving block x to y doesn't change the color of any of the blocks. Also the location of some other block z (different from x and y) is still the same. We could go on and on. To be able to derive that the color of any block isn't affected by the move, we would have to write something like:

$$\forall x, y, z, c, s. Holds(Color(z, c), s) \rightarrow Holds(Color(z, c), Result(Move(x, y), s)) \tag{6}$$

To say that other blocks don't move requires an axiom like:

$$\forall x, y, v, w, s. Holds(On(v, w), s) \wedge x \neq v \rightarrow Holds(On(v, w), Result(Move(x, y), s)) \tag{7}$$

We have to write similar axioms for every fluent that is not affected by an action. If there are n fluents and m actions, then we will have to write something on the order of $m \times n$ frame axioms! This is the essence of the frame problem. It's not that we can't represent change; we just can't represent it compactly.

It also seems non-intuitive that the fact that the fluents don't change must be represented. A person doesn't need to re-compute all unaffected fluents after an action takes place. A robot, if it is going to do efficient reasoning, shouldn't have to either.

One last note: the frame problem is not limited to situation calculus – it pops up in any logical representation which formalizes actions and change. The next few sections talk about possible solutions to the frame problem, and how they fail in one way or another.

1.3.1 Formalizing the Common Sense Law of Inertia

The common sense law of inertia states: "Normally, given any action (or event type) and any fluent, the action doesn't affect the fluent." We can say more formally: unless an action/fluent pair is abnormal in some way, in a situation, the fluent will have the same value both before and after the action takes place. In logic this becomes:

$$\forall a, f, s. \neg Ab(a, f, s) \rightarrow (Holds(f, s) \equiv Holds(f, Result(a, s))) \tag{8}$$

The equivalence above preserves both negative and positive information – if f was false in situation s, it will remain false in the situation $Result(a, s)$, assuming $\neg Ab(a, f, s)$.[5]

The idea of this approach is to combine the initial observations, like $Holds(On(A, Table), S0)$, with the effect axioms for each action. Then append equation 8 (instead of writing $m \times n$ frame axioms) to get a logical theory Σ. Circumscription can do the work of the $m \times n$ frame axioms, by minimizing the extent of Ab in the theory Σ, and letting $Holds$ vary.[6] (This is written $Circ[\Sigma; Ab; Holds]$.) Hopefully this theory $Circ[\Sigma; Ab; Holds]$ will have a first order equivalent, which will also dictate that only the fluents we want will persist.

Intuitively, this should work – by circumscribing Ab we guarantee that as many fluents' values as possible will persist ($Holds(f, s) \equiv Holds(f, Result(a, s))$) when the theory allows them to. By default, most fluents persist, which is the notion behind the common sense rule of inertia.

1.3.2 The Yale Shooting Problem

Unfortunately, in 1986, Hanks and McDermott found a counterexample which got unintuitive results. The counterexample is known as the "Yale Shooting Problem" [HM86], and is now the benchmark used for testing other solutions to the frame problem. The story is that there are three actions, *Load*, $Wait$, and $Shoot$, a gun, and a person, named Fred. There are two fluents: $Alive$ and *Loaded*. The gun is initially unloaded and Fred is alive. The object is to predict that Fred will be dead after loading the gun, waiting, and shooting.

In logic we can formalize the entire theory Σ as:

$$\begin{aligned}
\Sigma \equiv\ & \\
& Holds(Alive, S0) \\
& \neg Holds(Loaded, S0) \\
& \\
& \forall s. Holds(Loaded, Result(Load, s)) \\
& \forall s. Holds(Loaded, s) \rightarrow \neg Holds(Alive, Result(Shoot, s)) \\
& \forall s. \neg Holds(Loaded, Result(Shoot, s)) \\
& \\
& \forall a, f, s. \neg Ab(a, f, s) \rightarrow (Holds(f, s) \equiv Holds(f, Result(a, s)))
\end{aligned} \tag{9}$$

[5] We can rewrite $Holds(f, s) \equiv Holds(f, Result(a, s))$ in the equivalent form $Holds(f, s) \rightarrow Holds(f, Result(a, s)) \wedge \neg Holds(f, s) \rightarrow \neg Holds(f, Result(a, s))$.

[6] If you wonder why $Holds$ must vary, think about the minimization model-theoretically. When comparing models for Ab, you don't want to have to restrict yourself only to models which have the same extension for $Holds$. If anything you want to make your ordering relation $\leq_{Ab}$ as strong as possible to get the fewest minimal models.

The first two axioms above are the observation sentences. The next three are our effect axioms, and the last is the common sense law of inertia. What we would like to have is that:

$$Circ[\Sigma; Ab; Holds] \models \neg Holds(Alive, Result(Shoot, Result(Wait, Result(Load, S0)))) \quad (10)$$

Unfortunately this doesn't hold, because we get unintended models of $Circ[\Sigma; Ab; Holds]$.

Problem 4:[25 points] In order to get the result of the circumscription above, we could do some fancy syntactic manipulations. This is messy and not so intuitive. Another way is to enumerate and inspect the models of $Circ[\Sigma; Ab; Holds]$. It turns out each model can be represented as a graph with the fluents *Loaded* and *Alive* as boolean functions of situations, so not only are the models easy to determine, but simple to interpret.[7]

Below is a template graph – we know that *Loaded* is false in $S0$, and *Alive* is true. After the *Load* action, we know that the gun must be *Loaded*, according to the third axiom in 9. Finally, we also know that the gun is not *Loaded* after the *Shoot* action. These set values are blacked in. This leave us four intervals (dashed) over which the fluents can be true or false, yielding $2^4 = 16$ possible combinations.

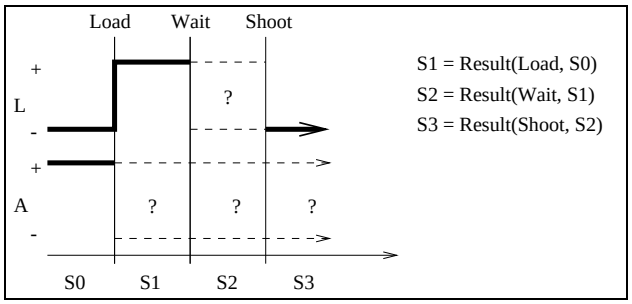

Figure 5: A Template Graph to Represent the Yale Shooting Problem.

Following is a graph corresponding to the intended model – the gun stays loaded after the *Wait* action, and Fred dies at the end of the *Shoot* action:

This model has three abnormalities, one for each discontinuity in the graph. Logically the *Ab*s are: $Ab(Load, Loaded, S0)$, $Ab(Shoot, Loaded, Result(Wait, Result(Load, S0)))$, and $Ab(Shoot, Alive, Result(Wait, Result(Load, S0)))$.

1. **[10 points] Enumerate the remaining eleven unintended models of $Circ[\Sigma; Ab; Holds]$, by trying the other fifteen possibilities for the unspecified intervals. (Four of the fifteen won't be legal because they won't obey the fourth axiom in 9.) Draw them all out in the same form as in figure 6. (Yes it is a pain but it is good for you!)**

2. **[10 points] By inspection, you should be able to draw a Hasse diagram based on the ordering $\leq_{Ab;Holds}$ on all twelve models. [Hint: Since *Holds* is varied,**

[7] Actually, each graph of the form in Figure 6 represents a whole class of models – we haven't said what happens in the situations "after" $S3$, or in the situations $Result(Wait, S0)$, etc. This is just a technicality but a major simplification that you should understand.

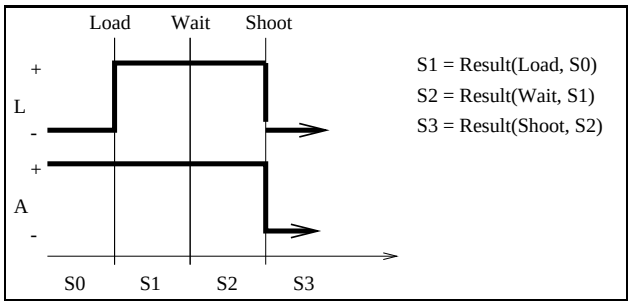

Figure 6: Intended Model of the Yale Shooting Problem.

you don't need to worry about the values of *Loaded* and *Alive*. All you have to do is order the models based on subset inclusion of the *Ab*s, which exactly correspond to transitions in the time functions.]

[**Hint #2: You can split the twelve models into two groups which are mutually incomparable if you look at the *Ab*s from the behavior of *Loaded* – four will have one *Loaded* function while eight will have another, and the *Ab*s in each are not inclusive of the other.**]

[**Hint #3: One set will look like a tree and the other a cube.**]

3. **[5 points] We can read off the minimal models by inspection of the Hasse diagram. It turns out there are four *Ab*-minimal models. For each, note how many *Abs* there are, and give a brief English description of each.**

1.3.3 Ensuring Fred Dies

The Yale Shooting Problem demonstrates that the naive approach for minimizing change does not work. Other solutions are mentioned in the book:

1. **Chronological minimization:** The unintended models for the Yale Shooting Problem all have one thing in common – change happens too soon. The idea behind chronological minimization is to prefer models which delay change as much as possible.

   For the Yale Shooting Problem, this approach will deliver the intended model where Fred dies. Unfortunately, there is no reason to prefer change to happen later rather than sooner. The "Stolen Car Scenario" [Kau86] is an example which highlights the arbitrariness of chronological minimization: Say a person parks his car. He waits one hour, and then another hour. When he returns to the parking space at 2pm he finds the car missing – it has been stolen! The car must have been stolen either during the first wait (from noon to 1pm) or the second wait (from 1pm to 2pm). Common sense dictates that either of these times are equally likely, but chronological minimization will declare the car was stolen between 1 and 2pm.

   Chronological minimization seems to be limited to *temporal projection problems*, that is, problems where you are given initial facts and effect axioms and then find out what happens in the future. It also seems to rely on the fact that the initial situation ($S0$) is completely known. The stolen car scenario is an example of an explanation problem – given a fact in the future, and initial facts, find out what happened in the interim to explain things.

Problem 5:[10 points] In chronological minimization, we look for the minimal $\leq_{chron}$ model, where $\leq_{chron}$ is defined as:

$M_1 \leq_{chron} M_2 \iff$

(a) M_1 **and** M_2 **agree on the interpretation of everything except** Ab **and** $Holds$**.**

(b) for each situation term σ_1**, either**

i. the Ab**s that occur at** σ_1 **in** M_1 **are a subset of those at** σ_1 **in** M_2

ii. there is an earlier situation term σ_2 **where the** Ab**s that occur at** σ_2 **in** M_1 **are a strict subset of those at** σ_2 **in** M_2

We can write this in math as:

$$\begin{aligned} M_1 \leq_{chron} M_2 \quad \iff & \\ & \forall S.(S \neq Ab \wedge S \neq Holds) \rightarrow M_1[\![S]\!] = M_2[\![S]\!] \wedge \\ & \forall \sigma_1 \in Situation.(Abextent(\sigma_1, M_1) \subseteq Abextent(\sigma_1, M_2)) \vee \\ & \quad \exists \sigma_2.\sigma_2 < \sigma_1 \wedge (Abextent(\sigma_2, M_1) \subset Abextent(\sigma_2, M_2)) \end{aligned} \tag{11}$$

where $Abextent$ **is defined as:**

$$Abextent(\sigma, M) = \{\langle a, f\rangle | \langle a, f, \sigma\rangle \in M[\![Ab]\!]\}$$

and $\sigma_1 < \sigma_2$ **just means that the situation** σ_2 **is a descendant of situation** σ_1**.**[8]

Given this definition of $\subseteq_{chron}$**, take the twelve models of** Σ **from Problem 4 and show that the intended model is chronologically smaller than the other 11 using the definition above.**

2. **Causal Minimization:** Introduced by [Hau87] and [Lif87], this is a more elegant approach, and also more intuitive. First two new predicates are added to the language, $Causes(a, f, v)$ (a causes f to take value v), and $Precond(f, v, a)$ (f has to have value v for a to succeed. (v can be either "true" or "false".) There are also two auxiliary predicates, $Success(a, s)$ and $Affects(a, f, s)$, which are defined in terms of $Causes$ and $Precond$. Instead of using the common sense law of inertia with Ab, we write axioms (not listed here) which say that

   (a) If action a causes fluent f to take on value v (true or false), and the action succeeds, then f will be v in the situation $Result(a, s)$.

   (b) If action a doesn't affect fluent f in s, then the value of f stays the same in $Result(a, s)$.

   (c) (definition of $Success$) a succeeds in s iff every fluent f which needs to have the value v in order for a to happen does so.

   (d) (definition of $Affects$) a affects f in situation s iff a succeeds in s and a causes f to take on some value (either true or false).

   We write a theory now in terms of the new predicates $Causes$ and $Precond$, and then take the theory $Circ[\Sigma; Causes, Precond; Holds]$. Circumscribing $Precond$ directly addresses

[8] We could define this as $\sigma_1 < \sigma_2 \iff (\exists a.\sigma_2 = Result(a, \sigma_1)) \vee (\exists a, s.\sigma_2 = Result(a, s) \wedge \sigma_1 < s)$.

the *qualification problem.*[9] Minimizing $Causes$ ensures that a fluent will change only if there is a cause for it to do so.

This works for the Yale Shooting Problem, and for all problems which are of the type where one predicts the future. The theory of causal minimization is well understood, and built on some nice principles which determine that it won't work for explanation problems, which are problems where we are given the beginning and end situations, and asked to find out what happened in between. It also cannot represent context-dependent effects (Toggling a switch is either "On" or "Off", depending on the previous state), non-deterministic actions, and theories which contain domain constraints. (At least not without some tinkering.)

Problem 6:[5 points] Write the four axioms needed for causal minimization in first-order logic, using only the predicates $Affects$, $Causes$, $Holds$, $Success$, $Precond$, and $=$, along with the truth value constants $true$ and $false$. Feel free to check your answer against the book.

3. **State Based Minimization**. This approach [Bak91], is similar to the original approach of taking an action theory and circumscribing Ab ($Circ[\Sigma; Ab; Holds]$). The main two differences, is that we let $Result$ vary instead of $Holds$, and add an axiom that says that a situation exists for every legal combination of fluents.

   So instead our grand theory looks like $Circ[\Sigma \wedge E; Ab; Result]$, where E is the existence of situations axiom. For the Yale Shooting Problem, it would be:

$$\begin{array}{l} \exists s.(Holds(Alive, s) \wedge Holds(Loaded, s)) \wedge \\ \exists s.(\neg Holds(Alive, s) \wedge Holds(Loaded, s)) \wedge \\ \exists s.(Holds(Alive, s) \wedge \neg Holds(Loaded, s)) \wedge \\ \exists s.(\neg Holds(Alive, s) \wedge \neg Holds(Loaded, s)) \end{array} \tag{12}$$

   State based minimization is a bit hard to understand, but it does solve the Yale Shooting Problem correctly.

   The motivation is that the Ab predicate is sensitive to its situation argument. When we say $Ab(a, f, s)$, we mean to identify the situation s by the fluents that are true in it (by its *state*), rather than the chain of actions that defines it ($s = Result(a_1, ...Result(a_n, S0)...)$). This is where we got the situation state interpretation of sitcalc.

   By ensuring that a situation exists for every combination of fluents, we can let the $Result$ function vary, letting it choose what states to map to, in order to minimize Ab.

   There is a lot more machinery that is needed for this technique to work, including extra domain closure axioms[10], prioritized circumscription, and compound fluents, which is best explained in Chapters 6 and 7 of [Sha97].

4. **Explanation Closure:** [Haa87] noted that axioms can be hand-crafted in a more compact form than frame axioms, so that we can remain in the realm of classical logic. For each fluent f, two axioms are written:

[9] The qualification problem is the fact that there are "an impractical and implausible number of qualifications" [McC80] which must be included to get all the preconditions for success of an action. Minimizing the extent of $Prec$ is the same as asserting that the only preconditions that hold for an action a are those we explicitly stated or that follow from a theory.

[10] The existence of situations axiom ensures that there are *enough* situations, domain closure is needed to make sure not too many exist.

$$\begin{array}{l}\forall a, s.(Holds(f, s) \wedge \neg Holds(f, Result(a, s))) \rightarrow \\ \quad ((a = action_1 \wedge (\text{whatever needs to } Hold \text{ in } s \text{ for } action_1 \text{ to happen})) \\ \quad \vee ... \vee \\ \quad (a = action_n \wedge (\text{whatever needs to } Hold \text{ in } s \text{ for } action_n \text{ to happen}))) \\ \\ \forall a, s.(\neg Holds(f, s) \wedge Holds(f, Result(a, s))) \rightarrow \\ \quad ((a = action'_1 \wedge (\text{whatever needs to } Hold \text{ in } s \text{ for } action'_1 \text{ to happen})) \\ \quad \vee ... \vee \\ \quad (a = action'_m \wedge (\text{whatever needs to } Hold \text{ in } s \text{ for } action'_m \text{ to happen})))\end{array} \tag{13}$$

(The $action_i$ above make f false, and the $action'_i$ above make f true.)

These explanation closure axioms give the *only* reason why a fluent changes from true to false, or false to true. These axioms, along with the effect axioms and observation sentences, form the desired theory.

Here are the explanation closure axioms for the Yale Shooting Problem:

$$\begin{array}{l}\forall a, s.(Holds(Loaded, s) \wedge \neg Holds(Loaded, Result(a, s))) \rightarrow (a = Shoot) \\ \\ \forall a, s.(\neg Holds(Loaded, s) \wedge Holds(Loaded, Result(a, s))) \rightarrow (a = Load) \\ \\ \forall a, s.(Holds(Alive, s) \wedge \neg Holds(Alive, Result(a, s))) \rightarrow (a = Shoot \wedge Holds(Loaded, s)) \\ \\ \forall a, s.(\neg Holds(Alive, s) \wedge Holds(Alive, Result(a, s))) \rightarrow \bot\end{array} \tag{14}$$

The fourth explanation closure looks strange – all it says is that there are no conditions ($\bot$) under which Fred is not $Alive$ in a situation and then $Alive$ in a resulting situation. This statement is equivalent to:

$$\forall a, s. \neg(\neg Holds(Alive, s) \wedge Holds(Alive, Result(a, s)))$$

or

$$\forall a, s. Holds(Alive, Result(a, s)) \rightarrow Holds(Alive, s)$$

5. **Successor State (a variation):** These axioms are also hand-crafted, and also do not require non-monotonic machinery. The basic form of these axioms is:

$$\begin{array}{l}\forall a, f, s. Holds(f, Result(a, s)) \equiv \\ \quad \bigvee_i \{a_i = \text{<action which makes f true>} \wedge \text{<all preconditions of } a_i \text{ hold>}\} \\ \quad \vee \\ \quad Holds(f, s) \wedge \neg \bigvee_j \{a_j = \text{<action which makes f false>} \wedge \text{<all preconditions of } a_j \text{ hold>}\}\end{array} \tag{15}$$

The intuition behind the logic is trivial: A fluent will be true after action a in exactly two cases. Either one of the actions which made it true succeeds, or it was true before and none of the actions that can falsify f *successfully* occur.

Note that there is one explanation closure axiom for every fluent f.

2 Otter

Otter is a theorem prover written by William McCune. It is a resolution-style theorem proving program that uses first-order logic with equality. It has been used to prove some complex theorems, including the fact that Robbins algebras are Boolean.
(See `http://www-unix.mcs.anl.gov/~mccune/papers/robbins/`) .
Otter comes with many bells and whistles, which are not necessary for this problem set. This section just shows you how to run Otter, and write logic in a form Otter can understand. The manual is available on the Leland system at `/usr/class/cs323/Otter-3.0.4/document/`. There's also a URL, over at `http://www-unix.mcs.anl.gov/AR/Otter`. You can download a free version for your PC if you end up enjoying theorem proving.

2.1 Commands

To start, first log onto a saga (I've tested it on these machines). From your home directory, try this command:

```
/usr/class/cs323/bin/Otter < /usr/class/cs323/examples/socrates.in
```

The output you get (try redirecting it to a file via >) should be the same as in `/usr/class/cs323/examples/socrates.out`. Make sure not to leave out the "<"!! Otter works by piping in files of logical axioms.

2.2 Format of Input File

The format of an Otter input file for this problem set is simple. The first statement is `set(auto)`, which tells the program to automatically choose a theorem proving strategy based on the structure of the theory. Then comes `formula_list(usable).`, which begins the lists of logical axioms. The list is ended by `end_of_list.`. Note that every command (and axiom) is punctuated with a period.

The axiom `all _x (_x = _x).` must be in your `formula_list`, otherwise equality won't be handled correctly. After that, you are free to write your own axioms. Here is a table of relevant logical operators:

negation	-
disjunction	|
conjunction	&
implication	->
equivalence	<->
existential quantification	exists
universal quantification	all

Predicates, constants, and functions may begin with either upper or lower case variables. Otter has some flaky rules on what is a variable, so it is recommended that all variables be explicitly preceded with a "_". (The underscore signals that it is definitely a variable!) If your theorems do not seem to work, please check that all variables have been under-scored!

The logical symbols `- | & -> <->` are used as usual. Precedence should be the same as well, though you should use parentheses if you are not sure. The form of axioms with quantification is different from convention, as shown in these examples:

```
all _a _b _c (P(_a) <-> (Q(_b, _c) & R(_c))).
exists _x (P(_x)).
```

There are no commas between quantified variables, and no period except at the end of the sentence. Also, the quantified part of the formula must be encapsulated in parentheses.

The conclusion, the formula you want to prove, also goes in the `formula_list`, except that the entire formula is negated. Otter works by resolving clauses until it hits $\perp$.

2.3 Format of Output File

The output file has eight sections:

1. A listing of the flags. (Ignore these!)
2. Copy of your formulas.
3. Your formulas converted to CNF.
4. A description of the clauses along with the strategy to be used.
5. Further processing of clauses
6. Resolution "Search"
7. Proof
8. Random Statistics

The only section that is important is (7) – the proof. (3) is a good place to check your axioms.

Problem 7:[10 points] "Everybody loves my baby, but my baby only loves me." Formalize this in first-order logic, and then use Otter to show that $Me = Baby$. Use the sample file at `/usr/class/cs323/examples/meandbaby.in` as an initial template. Please submit your file, and just the proof as your answer.

Problem 8:[10 points] Write the explanation closure axioms for the Yale Shooting Problem in Otter style. Combine with the initial observations and effect axioms (everything but the last axiom in equation 9). Use Otter to prove that

$$\neg Holds(Alive, Result(Shoot, Result(Wait, Result(Load, S0)))).$$

Use the sample file at `/usr/class/cs323/examples/ysp-expl.in` as an initial template.

Again, submit the file, and just the proof as your answer.

Problem 9:[10 points] Write the successor state axioms for the Yale Shooting Problem in Otter style. Combine with the initial observations. Use Otter to prove that $\neg Holds(Alive, Result(Shoot, Result(Wait, Result(Load, S0))))$. Use the sample file at `/usr/class/cs323/examples/ysp-succ.in` as an initial template.
I just want the file, and the proof.

References

[Bak91] A. B. Baker. Nonmonotonic reasoning in the framework of situation calculus. *Artificial Intelligence*, 49(1-3):5–23, 1991.

[Gal91] Antony Galton. **reified temporal theories and how to unreify them**[11]. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI'91)*, pages 1177–1182, 1991.

[GLR91] Michael Gelfond, Vladimir Lifschitz, and Arkady Rabinov. What are the limitations of the situation calculus? In Robert Boyer, editor, *Automated Reasoning: Essays in Honor of Woody Bledsoe*, pages 167–179. Kluwer Academic, Dordrecht, 1991.

[Haa87] A. R. Haas. The case for domain-specific frame axioms. *Proceedings of the 1987 Workshop on the Frame Problem*, pages 343–348, 1987.

[Hau87] Brian A. Haugh. Simple causal minimizations for temporal persistence and projection. In *In Proceedings of the Sith National Conference on Artificial Intelligence*, pages 218–223, 1987.

[HM86] S. Hanks and D. McDermott. Default reasoning, nonmonotonic logics and frame problem. In *Proceedings of AAAI-86*, pages 328–333. Morgan Kaufmann, 1986.

[Kau86] H. Kautz. The logic of persistence. *Proceedings of the AAAI 86*, pages 401–405, 1986.

[Lif87] Vladimir Lifschitz. Formal theories of action. In F. Brown, editor, *The Frame Problem in Artificial Intelligence: Proceedings of the 1987 Workshop*, pages 35–58, Los Altos, CA, 1987. Morgan Kaufmann Publishers.

[McC80] John McCarthy. Circumscription—A Form of Non-Monotonic Reasoning[12]. *Artificial Intelligence*, 13:27–39, 1980. Reprinted in [McC90].

[McC90] John McCarthy. *Formalizing Common Sense: Papers by John McCarthy.* Ablex Publishing Corporation, 355 Chestnut Street, Norwood, NJ 07648, 1990.

[MH69] John McCarthy and Patrick J. Hayes. Some Philosophical Problems from the Standpoint of Artificial Intelligence[13]. In B. Meltzer and D. Michie, editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, 1969.

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

[11] http://www.dcs.ex.ac.uk/ antony/abstracts/ijcai91.html
[12] http://www-formal.stanford.edu/jmc/circumscription.html
[13] http://www-formal.stanford.edu/jmc/mcchay69.html