CS323 - Winter 97 - Homework #2 – Solutions

February 12, 1997

1 The Hanks-McDermott shooting problem

Definition 1.1 *The* Hanks-McDermott shooting problem *(sometimes refered to as the "Yale shooting Problem") is the following story: A person can be either* ALIVE *or* DEAD. *A gun can be either* LOADED *or* UNLOADED. *At a known situation the person is alive, and a gun become loaded whenever a* LOAD *action occurs. Any time a person is shot (i.e., a* SHOOT *action) with a loaded gun, he becomes dead.*

Problem 1.2 1. *Formalize the problem in Situation Calculus. Emphasize which are your Frame Axioms.*

2. *Show the process of LOAD;SHOOT as a chain of situations, emphasising what sentences are true in each situation.*

3. *Add the action WAIT (with the apropriate axioms) to your Database. Show the process of LOAD;WAIT;SHOOT as a chain of situations, emphasising what sentences are true in each situation.*

SOLUTION

1. We take the language of Situation Calculus to be a 3-sorted language: $Fluents$, $Actions$, and $Situations$. We will denote variables for the three sorts by f, a, s, respectively, and subscripted versions thereof. Our language includes the object constants $S0$ ($Situations$), $LOAD$ and $SHOOT$ ($Actions$), and $ALIVE$ and $DEAD$ ($Fluents$). It also includes the function constant $Result$ ($Situations$, $Actions \rightarrow Situations$), and the relations constant $Holds$ ($Fluents$, $Situations$).

   Our axioms will be as follows:

   $Holds(ALIVE, S0) \wedge \neg Holds(LOADED, S0)$

   $\forall s\ Holds(LOADED, Result(LOAD, s))$
   $\forall s\ Holds(LOADED, s) \implies \neg Holds(ALIVE, Result(SHOOT, s))$

   Frame Axioms:
   $\forall s\ \forall a \neq SHOOT\ Holds(ALIVE, Result(a, s)) \iff Holds(ALIVE, s)$
   $\forall s\ \forall a \neq LOAD\ Holds(LOADED, Result(a, s)) \iff Holds(LOADED, s)$

2. Let $S1 = Result(LOAD, S0)$, $S2 = Result(SHOOT, S1)$. Then,

$$\begin{array}{l} Holds(ALIVE, S0) \wedge \neg Holds(LOADED, S0) \\ Holds(ALIVE, S1) \wedge Holds(LOADED, S1) \\ \neg Holds(ALIVE, S2) \wedge Holds(LOADED, S2) \end{array}$$

3. We add the constant $WAIT$ of sort $Actions$. Let $S1 = Result(LOAD, S0)$, $S2 = Result(WAIT, S1)$, $S3 = Result(SHOOT, S2)$. Notice that our previous frame axioms suffice for our purposes. Then,

$$\begin{array}{l} Holds(ALIVE, S0) \wedge \neg Holds(LOADED, S0) \\ Holds(ALIVE, S1) \wedge Holds(LOADED, S1) \\ Holds(ALIVE, S2) \wedge Holds(LOADED, S2) \\ \neg Holds(ALIVE, S3) \wedge Holds(LOADED, S3) \end{array}$$

2 The Missionaries and Cannibals Problem

In this Problem Set we will deal (again) with a *soft* version of the MCP we will call the Elementary-MCP.

Definition 2.1 *The* Elementary-K Missionaries and Cannibals Problem *is the following problem:*

> *K missionaries and K cannibals get to a river. There is a bridge over the river, and they want to cross over to the other bank. How shall they cross?*

Problem 2.2 *1. Formalize the Elementary-2-MCP in Situation Calculus. Alter the formalism given for the Elementary-1-MCP in order to run it on the PVS. Prove that your formalism is solvable using the PVS. Point out the Frame Axioms you use.*

SOLUTION - on the web.

3 Circumscription

Problem 3.1 *Circumscribe the following formulas (these should become simple 1st order formulas describing the circumscribed predicate).*

1. $CIRC[isblock(A) \vee isblock(B); isblock]$
2. $CIRC[isblock(A) \vee isblock(B); isblock; A, B]$
3. $CIRC[\forall x \neg P(x); P]$
4. *Let x be a tuple of variables, $F(x)$ a formula with all parameters explicitly shown. Prove that if $F(x)$ does not contain P, then*

$$CIRC[\forall x(F(x) \implies P(x)); P] \equiv \forall x(F(x) \equiv P(x))$$

(i.e., in some cases Circumscription *is reducible to* predicate completion*).*

5. *Show that* $Circ[Bird(Tweety) \wedge (\exists x)(Bird(x) \wedge \neg Fly(x)); \neg Fly; Bird, Fly]$ *does not imply* $Fly(Tweety)$.

Bonus *Prove that for a predicate* P *there is no 1st order formula formulating the* transitive closure *of* P.

SOLUTION Here I will detail only those proofs that went wrong for some of you. For the rest, I will write only the final solution.

1. $Circ[isblock(A) \vee isblock(B); isblock] \equiv isblock = \{A, B\}$.

2. $CIRC[isblock(A) \vee isblock(B); isblock; A, B]$ is (expanding the circumscription formula)

$$isblock(A) \vee isblcok(B) \wedge \forall p, a, b\, [(p(a) \vee p(b)) \wedge p \leq isblock \implies p = isblock]$$

Let $p = \{A\}$, $b = a = A$. Then, the inner formula becomes

$$p(A) \vee p(A) \wedge p \leq isblock \implies p = isblock$$

which is equivalent (because of our choice of p) to $isblock(A) \implies isblock = \{A\}$. We do the same for $p = \{B\}$ and $b = a = B$, and we get $isblock(B) \implies isblock = \{B\}$. Thus $isblock = \{A\} \vee isblock = \{B\}$. ■

3. $CIRC[\forall x \neg P(x); P] \equiv \forall x \neg P(x)$

4. $CIRC[\forall x(F(x) \implies P(x)); P]$ is equal to (expanding the circumscription formula)

$$\forall x(F(x) \implies P(x)) \wedge \forall p\, [(\forall x(F(x) \implies p(x))) \wedge p \leq P \implies p = P]$$

Take $p = \lambda x\ F(x)$, and get that $P = \lambda x\ F(x)$. The "trick" is in choosing p to be this way. We can do that because F does not contain P[1]. In that case, we are sure that when we choose p, $F(x)$ is fixed (nothing but P varies), and thus chosing p this way, is well defined. ■

5. Let $\mathcal{M}$ be an interpretation with the world $U = \{Tweety\}$, and the assignments $Bird = \{Tweety\}$, and $Fly = \emptyset$. $\mathcal{M} \models Bird(Tweety) \wedge (\exists x)(Bird(x) \wedge \neg Fly(x))$, but it is obvious that we can not make $\neg Fly$ smaller, as the only element in our model is $Tweety$, and thus it must remain a $Bird$ (and thus must not Fly). So $\mathcal{M} \models Circ[Bird(Tweety) \wedge (\exists x)(Bird(x) \wedge \neg Fly(x)); \neg Fly; Bird, Fly]$, but $\mathcal{M} \models \neg Fly(Tweety)$. ■

[1] If F contains P then we can not be sure that there is a minimal element (i.e., the circumscription might not be consistent). Example: take $F(x)$ to be $\neg P(x)$.

Bonus Assume that for every binary relation P there is a 1st order formula φ_P that gives the transitive closure of P. Let P be the successor relation $P(s, s+1)$. If we have a transitive closure of P, then P is the set of Natural numbers, and we can write $\forall x P(x)$, representing that the only models possible are countable. Contradicting to the well known theorem that if a theory has a countable model, then for every infinite cardinal, that theory has a model of that cardinal.[2] ■

[2]This is the shortest proof that came to my mind, but I think there are some others.