CS 323, Fall 2000
Problem Set #1: Solutions to "Formalizing Commonsense Knowledge and Reasoning: An Introduction"
Handout #7

These are the answers to the problems posed in "Formalizing Commonsense Knowledge and Reasoning: An Introduction." This paper can be found at:
`http://www-formal.Stanford.EDU/pub2/papers/form.ps`.

$$Ontable(A), Ontable(B), On(C, A) \tag{1}$$

$$\forall x. x = A \lor x = B \lor x = C \tag{2}$$

$$\forall x, y, z. On(x, y) \land On(x, z) \rightarrow y = z \tag{3}$$

$$\forall x, y. On(x, y) \rightarrow \neg Ontable(x) \tag{4}$$

$$A \neq B, B \neq C, C \neq A \tag{5}$$

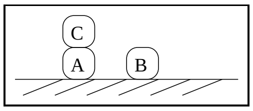

Figure 1: What T^b looks like.

T^b = Axioms (1)–(5)

$$M \models T^b$$
$$|M| = \{A^*, B^*, C^*\}, M[\![A]\!] = A^*, M[\![B]\!] = B^*, M[\![C]\!] = C^*$$
$$M[\![On]\!] = \{\langle C^*, A^*\rangle\}, M[\![Ontable]\!] = \{A^*, B^*\}$$

1. **[10 points] Prove that every model of T^b is isomorphic to M (i.e. for every model M' of T^b there exists a 1–1 correspondence between the universe of M and the universe of M' which maps the interpretation of any symbol s in M into its interpretation in M').**

   Plan: Pick any model M' of T^b, and then construct an isomorphism $i(x)$ from M to M'. Remember an isomorphism is a homomorphism which is 1–1. A homomorphism is a function $h : |M| \rightarrow |M'|$ where:

(a) For each relational symbol R:

$$\langle a_1, ..., a_n \rangle \in M[\![R]\!] \iff \langle h(a_1), ..., h(a_n) \rangle \in M'[\![R]\!]$$

(b) For each function symbol f:

$$h(M[\![f]\!](a_1, ..., a_n)) = M'[\![f]\!](h(a_1), ..., h(a_n))$$

(c) For each constant symbol c:

$$h(M[\![c]\!]) = M'[\![c]\!]$$

Without loss of generality we can name the constants in T^b:

$$M'[\![A]\!] = a, M'[\![B]\!] = b, M'[\![C]\!] = c$$

At the very least, $|M'| = \{a, b, c, ...\}$. Let's show that $|M'| = \{a, b, c\}$. Say there was another element $x \in |M'|$, different from a, b, c. Thus $x \neq a, x \neq b$, and $x \neq c$. But axiom (2) asserts the opposite. Hence $|M'| = \{a, b, c\}$.

By (5), the elements are all distinct, so M' has exactly 3 distinct elements. We can construct our isomorphism on the object constants, by asserting $i(A^*) = a, i(B^*) = b$, and $i(C^*) = c$. We can verify that this is 1–1. Now we just need to extend $i(x)$ to the relational constants On and $Ontable$.

Since $M' \models Ontable(A) \wedge Ontable(B) \wedge On(C, A)$, we have $\{a, b\} \subseteq M'[\![Ontable]\!]$, and $\{\langle c, a \rangle\} \subseteq M'[\![On]\!]$. To prove the isomorphism on the relational constants, we need to show that $x \in M[\![Ontable]\!] \iff i(x) \in M'[\![Ontable]\!]$, and $\langle x, y \rangle \in M[\![On]\!] \iff \langle i(x), i(y) \rangle \in M'[\![On]\!]$.

If we could show that $\{a, b\} = M'[\![Ontable]\!]$, and $\{\langle c, a \rangle\} = M'[\![On]\!]$ then we would be done. So let's assume there's an element $x \in M'[\![Ontable]\!]$, and $x \notin \{a, b\}$. By (2), $x = c$, but $On(C, A)$, so $\neg Ontable(C)$ by (4), so $c \notin M'[\![Ontable]\!]$. We've reached a contradiction, so $M'[\![Ontable]\!] = \{a, b\}$.

Let's do the same for $M'[\![On]\!]$. Say $\langle x, y \rangle \in M'[\![On]\!]$, where $\langle x, y \rangle \neq \langle c, a \rangle$. By axiom (4), since $On(x, y)$, $\neg Ontable(x)$. By the argument above we see that x could only be c. But then by axiom (3) $y = a$.

We've shown an isomorphism between any two models of T^b. Thus T^b is categorical.

Point Breakdown:
-1: Missing clarity.
-4: Talking about why other elements are excluded from $M[\![On]\!], M[\![Ontable]\!]$, but not showing why.
-5: Not showing why other elements are excluded from $M[\![On]\!], M[\![Ontable]\!]$.
-8: Improper proof.

2. **[10 points] Show that T^b admits quantifier elimination. Which axioms of T^b did you use in the proof?**

To show T^b admits quantifier elimination, we need to show that for any formula $\phi(\bar{x})$, there is another quantifier-free formula, $\phi^*(\bar{x})$ such that:

$$T^b \models \phi(\bar{x}) \equiv \phi^*(\bar{x})$$

First, we know that we can write any formula in prenex form, that is, with all quantifiers on the left side of the formula. We define the depth of the quantification to be the number of quantifiers on the outside.

We prove $T^b \models \phi(\bar{x}) \equiv \phi^*(\bar{x})$ by induction on the quantification depth of ϕ.

- Base Case (n = 0): $\phi(\bar{x})$ has no quantifiers. So, we are done. (It is equivalent to itself.)
- Inductive Step: Say $\phi(\bar{x})$ is of depth $n+1$. Then either:
  (a) $\phi(\bar{x}) \equiv \forall y.\psi(y, \bar{x})$, or
  (b) $\phi(\bar{x}) \equiv \exists y.\psi(y, \bar{x})$

where $\psi(y, \bar{x})$ is of quantification depth n, and by the inductive hypothesis is equivalent to a quantifier-free formula. And we can always write $\forall y.\psi(y, \bar{x})$ as $\psi(A, \bar{x}) \wedge \psi(B, \bar{x}) \wedge \psi(C, \bar{x})$, and $\exists y.\psi(y, \bar{x})$ as $\psi(A, \bar{x}) \vee \psi(B, \bar{x}) \vee \psi(C, \bar{x})$. These rewrites are quantifier-free so we are done.

The only axiom we used in this proof is axiom (2).

Point Breakdown:
-1: using $\vee$ instead of $\wedge$.
-1: leaving out variables in atomic formulas.
-1: Saying only atomic formulas are those in axiom (2).
-1: Not explicitly mentioning the domain closure axiom somewhere.
-2: No real argument, just equivalences of $\forall x.\psi(x)$, $\exists x.\psi(x)$.

3. [10 points] **Prove the completeness of T^b using quantifier elimination.**

We know any formula can be written modulo T^b as a quantifier-free formula. Hence any sentence (= closed formula) is equivalent to a quantifier-free sentence. But this is just a boolean combination of the atomic formulas of the form $On(t_1, t_2), Ontable(t_1)$, and $t_1 = t_2$, where $t_1, t_2 \in \{A, B, C\}$. By inspection we can determine the truth or falsity of any of these formulas in T^b:

Atom	Truth Value	Why?	Atom	Truth Value	Why?
$Ontable(A)$	T	Axiom (1)	$On(A, A)$	F	Axioms (1, 4)
$Ontable(B)$	T	Axiom (1)	$On(A, B)$	F	Axioms (1, 4)
$Ontable(C)$	F	Axioms (1, 4)	$On(A, C)$	F	Axioms (1, 4)
$A = A$	T	Reflexivity	$On(B, A)$	F	Axioms (1, 4)
$A = B$	F	Axiom (5)	$On(B, B)$	F	Axioms (1, 4)
$A = C$	F	Axiom (5)	$On(B, C)$	F	Axioms (1, 4)
$B = A$	F	Axiom (5)	$On(C, A)$	T	Axiom (1)
$B = B$	T	Reflexivity	$On(C, B)$	F	Axioms (1, 3, 5)
$B = C$	F	Axiom (5)	$On(C, C)$	F	Axioms (1, 3, 5)
$C = A$	F	Axiom (5)			
$C = B$	F	Axiom (5)			
$C = C$	T	Reflexivity			

And since every sentence of T^b is written as a boolean combination of these atoms, we can determine its truth or falsity as well.

Thus T^b is complete.

4. [**5 points**] **Show that $Clear(B)$ cannot be proved without axiom (5).**

   Let us assume the opposite. Call T^b_{-5} the theory without axiom (5), and say that $T^b_{-5}\vdash Clear(B)$. This means that $T^b_{-5}\cup\{\neg Clear(B)\}\vdash\perp$, or that $T^b_{-5}\cup\{\neg Clear(B)\}$ has no model. But consider the structure M', which is like M except $A = B$:

$$|M'| = \{A^*, C^*\}, M'[\![A]\!] = A^*, M'[\![B]\!] = A^*, M'[\![C]\!] = C^*$$
$$M'[\![On]\!] = \{\langle C^*, A^*\rangle\}, M'[\![Ontable]\!] = \{A^*\}$$

   $M \models T^b_{-5}\cup\neg Clear(B)$, which provides the contradiction.

   Point Breakdown:
   -1: confusion, especially about a formula being true in *all* interpretations versus some of them.
   -0: Many people just provided the other model without explaining really how this proves the question. You got credit, but please make you sure understand why this method works!

5. [**5 points**] **Show that axiom (2) cannot be derived from the other axioms.**

   Let us again assume the opposite and derive a contradiction. Call T^b_{-2} the theory without axiom (2), and assume that $T^b_{-2}\vdash(2)$. This is equivalent to $T^b_{-2}\cup\{\neg\forall x.x = A \vee x = B \vee x = C\}\vdash\perp$, or that $T^b_{-2}\cup\{\exists x.x \neq A \wedge x \neq B \wedge x \neq C\}$ has no model. But we can construct the following structure M' which is like M except with an extra element D^*:

$$|M'| = \{A^*, B^*, C^*, D^*\}, M'[\![A]\!] = A^*, M'[\![B]\!] = B^*, M'[\![C]\!] = C^*$$
$$M'[\![On]\!] = \{\langle C^*, A^*\rangle\}, M'[\![Ontable]\!] = \{A^*, B^*\}$$

   We can verify that $M' \models T^b_{-2}\cup\{\exists x.x \neq A \wedge x \neq B \wedge x \neq C\}$.

   Point Breakdown:
   -1: handwaving, or no clarity.
   -0: Many people just provided the other model without explaining really how this proves the question. You got credit, but please make you sure understand why this method works!

6. [**10 points**] **U_N is extended by the unary predicate *small* and the axioms**

$$\exists x.small(x), \exists x.\neg small(x)$$

   Prove that this extension is conservative if and only if $N > 1$.

   Let's call $U_N^+ = U_N\cup\{\exists x.small(x), \exists x.\neg small(x)\}$.

   U_N^+ is conservative over U_N if every theorem of U_N^+ which lacks *small* is provable in U_N.

   $\rightarrow$: One of the theorems of U_N^+ is that there are at least two distinct objects. (From: $\exists x.small(x) \wedge \exists x.\neg small(x)$). So $U_N^+\vdash\exists x, y.x \neq y$. Hence this theorem must be provable in U_N, which is only possible if $N > 1$.

   $\leftarrow$: Let's assume $N > 1$. We want to show that U_N^+ is a conservative extension of U_N, or that any theorem T of U_N^+ lacking the symbol *small* is true in U_N. (In math this is $U_N^+ \models T\rightarrow U_N \models T$.)

So, pick a T such that $U_N^+ \models T$. We need to show that $U_N \models T$, or that every model of U_N is also one of T. So pick a model M of U_N. All we have to do is show that $M \models T$.

Given any such model M of U_N, we can extend it to create another model $e(M)$ of U_N^+, where $e(M)$ has the same domain as M ($|e(M)| = |M|$), and the extension of *small* in $e(M)$ has at least one element of $e(M)$'s domain but not all ($\phi \subset e(M)[\![small]\!] \subset |e(M)|$). We can do this because there are at least two elements in the universe of any model of U_N/U_N^+ ($N > 1$).

Now $e(M) \models U_N^+$, and thus $e(M) \models T$, as T is a theorem of U_N^+. But in verifying the truth of T, we do not access the part of $e(M)$ dealing with *small*, since T lacks that symbol. But then $M \models T$ too, then, because M and $e(M)$ are the same everywhere else.

Point Breakdown:
-1: Handwaving over why if $U_N^+ \models \phi \rightarrow U_N \models \phi$, $\phi \in \mathcal{L}[U_N]$.
-1: Not saying how $\exists x.\neg small x$ is satisfied in the extended model.
-3: Some confusions.
-4: Saying that U_1^+ is contradictory.
-4, -5: Confusion over the theorems of U_N.
-6: No understanding of U_N, "When $N = 1$, U_N^+ is contradictory" + hand-wavy explanation.
-7: Above with more confusion.
-8: Yet even more confusion.

There was a lot of confusion about what U_N means. It means that there are at *least* N objects, not that there are exactly N objects!

7. **[5 points] Is U_{N+1} conservative over U_N?**

For U_{N+1} to be conservative, every theorem of U_{N+1} in the language of U_N must be provable in U_N. But consider the following theorem of U_{N+1}:

$$\exists x.x \neq C_1 \wedge ... \wedge x \neq C_N$$

This sentence, which asserts that there is an N+1th element different from $C_1, ..., C_N$, cannot be provable in U_N, as U_N has some models where there are only N elements.

So U_{N+1} is not conservative over U_N.

Point Breakdown:
-4: If you're wrong.

8. **[10 points] Define $Clear(t)$ as $\neg\exists b.(location(b) = t)$ (where t is a location term) and prove**

$$\neg Clear(top(A)), Clear(top(B)), Clear(top(C))$$

in T^{bl}.

T^{bl}	$\vdash location(C) = top(A)$	1. Given.
	$\vdash \exists b.location(b) = top(A)$	2. Existential Generalization
	$\vdash \neg Clear(top(A))$	3. Definition of *Clear*

The other two are easier to prove by contradiction. Assume $\neg Clear(top(B))$, or $\exists b.location(b) = top(B)$. Either $b = A$ or $b = B$ or $b = C$:

(a) $b = A$: $location(A) = top(B)$. Then $Table = top(B)$, which is a contradiction since $\forall b.top(b) \neq Table$.

(b) $b = B$: $location(B) = top(B)$. Once again we get $Table = top(B)$, which is a contradiction.

(c) $b = C$. $location(C) = top(B)$. Then $top(A) = top(B)$, and thus $A = B$, which is another contradiction.

Since $T^{bl} \cup \{\neg Clear(top(B))\} \vdash \perp$, we know $T^{bl} \vdash Clear(top(B))$.

Finally, let's prove $T^{bl} \vdash Clear(top(C))$. We can follow the steps as above to assume $\neg Clear(top(C))$ and thus that $\exists b.location(b) = top(C)$. Let's go through the cases again:

(a) $b = A$: $location(A) = top(C)$. Then $Table = top(C)$, which is a contradiction since $\forall b.top(b) \neq Table$.

(b) $b = B$: $location(B) = top(C)$. Once again we get $Table = top(C)$, which is a contradiction.

(c) $b = C$. $location(C) = top(C)$. Then $top(A) = top(C)$, and thus $A = C$, which is another contradiction.

Point Breakdown:
-1: skipping a few steps.
-2: skipping too many steps, without explanation.
-5: Handwaving at the proof.

9. **[10 points] Construct a model of T^{bl}. (To define its universe, you have to specify two disjoint sets, the domain $|M|^b$ of blocks and the domain $|M|^l$ of locations.)**

$$M \models T^{bl}$$

$$\begin{array}{l}
|M|^b = \{A^*, B^*, C^*\}, |M|^l = \{Table^*, t_a^*, t_b^*, t_c^*\} \\
M[\![A]\!] = A^*, M[\![B]\!] = B^*, M[\![C]\!] = C^*, M[\![Table]\!] = Table^* \\
M[\![top]\!] : |M|^b \to |M|^l \\
M[\![top]\!](A^*) = t_a^* \\
M[\![top]\!](B^*) = t_b^* \\
M[\![top]\!](C^*) = t_c^*
\end{array}$$

$$\begin{array}{l}
M[\![location]\!] : |M|^b \to |M|^l \\
M[\![location]\!](A^*) = Table^* \\
M[\![location]\!](B^*) = Table^* \\
M[\![location]\!](C^*) = t_a^*
\end{array}$$

You can verify that each axiom in T^{bl} is satisfied. We assume that all of the object constants are distinct.

Point Breakdown:
about -1: for each error.

10. **[10 points] Let ϕ be the assertion that there are locations other than the table and the tops of blocks. Show that ϕ can be neither proved nor refuted in T^{bl}.**

First,

$$\phi \equiv \exists l. l \neq Table \wedge \neg \exists b. top(b) = l.$$

To show that ϕ can neither be proved nor refuted, we need to construct a model for $T^{bl} \cup \{\phi\}$, and a model for $T^{bl} \cup \{\neg\phi\}$.

The solution above for Problem 9 is a model of $T^{bl} \cup \{\neg\phi\} = T^{bl} \cup \{\forall l. l = Table \vee \exists b. top(b) = l$. Let's just construct a model M' for $T^{bl} \cup \{\phi\}$. The idea is to just add a new location $Floor^*$:

$$\begin{array}{l}
|M'|^b = \{A^*, B^*, C^*\}, |M'|^l = \{Table^*, t_a^*, t_b^*, t_c^*, Floor^*\} \\
M'[\![A]\!] = A^*, M'[\![B]\!] = B^*, M'[\![C]\!] = C^*, M'[\![Table]\!] = Table^* \\
M'[\![top]\!] : |M'|^b \rightarrow |M'|^l \\
M'[\![top]\!](A^*) = t_a^* \\
M'[\![top]\!](B^*) = t_b^* \\
M'[\![top]\!](C^*) = t_c^* \\
\\
M'[\![location]\!] : |M'|^b \rightarrow |M'|^l \\
M'[\![location]\!](A^*) = Table^* \\
M'[\![location]\!](B^*) = Table^* \\
M'[\![location]\!](C^*) = t_a^*
\end{array}$$

You can verify that $M' \models T^{bl} \cup \{\phi\}$, again assuming that all object constants above are distinct.

Point Breakdown:
-2, -3: handwaving about there being no axioms that can either prove or refute ϕ.

11. **[10 points] Are there formulas without location variables which can be neither proved nor refuted in T^{bl}?**

The intuition is that the axioms force any model of T^{bl} to have exactly three block elements, A, B, C. The functions *top* and *location* add at most 6 location objects to the universe, and $Table$ is the 7th.

If we cannot quantify over locations, then these 7 terms are really the only ones we can use. We can quantify over blocks, but this is equivalent to $\wedge$-ing or $\vee$-ing over A, B, C.

The only relation in T^{bl} is $=$. So the only questions we can ask is whether two elements are equal or not. But the formalization provides enough information to find this out, for any two terms above.

Hence every formula's truth should be provable, if it lacks location variables.

One hard core, but clean way to prove this false is to show that T^{bl} is complete over Φ, where Φ is the set of formulas without location variables. We can show completeness, by proving every formula in Φ is equivalent to a quantifier-free formula without location variables, and then showing that every atomic formula is provable in T^{bl}.

We can copy the steps in Problems 2 and 3.

Again we assume that every formula in Φ is in prenex form, and show $T^{bl} \models \phi(\bar{x}) \equiv \phi^*(\bar{x})$ by induction on the quantification depth.

- Base Case (n = 0): Done.
- Inductive step: Say $\phi(\bar{x}) \equiv Qy.\psi(y,\bar{x})$. Since we are not allowed to have location variables, y must be a block variable. Then we can set $Qy.\psi(y,\bar{x}) \equiv \psi(A,\bar{x}) \cdot \psi(B,\bar{x}) \cdot \psi(C,\bar{x})$, where $\cdot = \wedge$ if $Q = \forall$, and $\cdot = \vee$ if $Q = \exists$.

Now we just need to show that every atomic formula is provable in T^{bl}, that is the set of equations $t_1 = t_2$, where $t_1, t_2 \in \{A, B, C, Table, location(A), location(B), location(C), top(A), top(B), top(C)\}$, and the sorts match up correctly. We can verify that this is possible.

Hence, there are no formulas without location variables that are not provable in T^{bl}.

Point Breakdown:
-1: leaving out the case when atomic formulas have a variable. (If proving by induction/quantifier elimination)
-1: No use of domain closure axiom.
-2: leaving out some atomic formulas.
-2: Not enough details.
-5: No justification for answer.

12. [**5 points**] **Can it happen that $\alpha(F)$ is provable in T^{bl} even though F is not provable in T^b?**

No. T^b is complete, so such F cannot exist.

Point Breakdown:
-2: lots of confusion.