1 1999 Paper 6 Question 10

https://www.cl.cam.ac.uk/teaching/exams/pastpapers/y1999p6q10.pdf

(a) Describe the role of Herbrand models in mechanical theorem proving. What may we infer when a set of clauses has no Herbrand model?

A Herbrand Model is an interpretation such that the clauses hold and the domain of the interpretation is the Herbrand Universe. Formally, a Herbrand Model is a pair $(\mathcal{H}, I)$ where $\mathcal{H}$ is the domain of the variables and I is a mapping from variables and function literals to elements in the Herbrand Universe $\mathcal{H}$ or concrete functions.

The goal of mechanical theorem provers (such as those using DPLL) is to *find* a Herbrand Model. If a set of clauses has no Herbrand Model; the set of clauses is *unsatisfiable.*

most general
finite refutations
syntactic universe

(b) Convert the following problem into clause form. Justify each step you take and explain in what respect the set of clauses is equivalent to the original problem.

$$\exists x\,[P(x) \land Q(x)] \longrightarrow \exists x\,[P(f(x,x)) \lor \forall y Q(y)]$$

To convert to clause form; we must Skolemize, then negate and convert to CNF. The clauses are the disjunctions.

whoa! Nooooo. Negate first. Skolemisation does not preserve meaning so if you negate you won't have the negation of the original formula!

For clarity, I will rename the variables such that the new expression is α-equivalent and has no duplicate names

$$\exists x\,[P(x) \land Q(x)] \longrightarrow \exists z\,[P(f(z,z)) \lor \forall y Q(y)]$$

Skolemize. We must firstly replace existentially bound variables with Skolem Constants or Skolem Functions as appropriate. In this case, both x and z become Skolem Constants

$$P(a) \land Q(a) \longrightarrow P(f(b,b)) \lor \forall y Q(y)$$

no

Now remove all universal quantifiers

$$P(a) \land Q(a) \longrightarrow P(f(b,b)) \lor Q(y)$$

no

Negate. Then use the equivalence $\neg(A \to B) \simeq A \land \neg B$

$$\begin{aligned} &P(a) \land Q(a) \land \neg(P(f(b,b)) \lor Q(y)) \\ \simeq &P(a) \land Q(a) \land \neg P(f(b,b)) \land \neg Q(y)) \end{aligned}$$

no

Now convert CNF to clauses by taking each disjunct as a clause

$$\{P(a)\}\ \{Q(a)\}\ \{\neg P(f(b,b))\}\ \{\neg Q(y))\}$$

make it stop

Skolemization does not preserve meaning or validity of a formula. However, it retains inconsistency. The final set of clauses is inconsistent if and only if the original formula was inconsistent.

(c) Describe the Herbrand universe for your clauses.

The Herbrand universe $\mathcal{H}$ for these clauses is described as follows:

$$\begin{aligned} H_0 &= \{c, d\} \\ H_{i+1} &= H_i \cup \{f(a,b) \mid a, b \in H_i\} \\ \mathcal{H} &= \bigcup_{i \geq 0} H_i \end{aligned}$$

Harry Langford
hjel2@cam.ac.uk

The Herbrand Universe can also be defined recursively as follows:

$$\mathcal{H} = \{f(x, y) \mid x, y \in \mathcal{H}\}$$

(d) Produce a resolution proof from your clauses or give reasons why none exists.

To prove the theorem, we must derive the empty clause $\square$

$$\{P(a)\}\ \{Q(a)\}\ \{\neg P(f(b, b))\}\ \{\neg Q(y))\}$$

Firstly, resolve $Q(a)$ with $\neg Q(y)$ using the unifier $\sigma = [a/y]$

$$\begin{aligned} &\sim\{P(a)\}\ \{Q(a)\}\ \{\neg P(f(b, b))\}\ \{\neg Q(a))\} \\ &\sim\{P(a)\}\ \{\neg P(f(b, b))\}\ \square \end{aligned}$$

The empty clause been derived so the original theorem is proved

(e) Exhibit a Herbrand model for your clauses or give reasons why none exists.

The clauses are unsatisfiable – therefore they have no Herbrand Model.

2 1998 Paper 5 Question 10

https://www.cl.cam.ac.uk/teaching/exams/pastpapers/y1998p5q10.pdf

(b) Attempt to prove the above formula using the sequent calculus until either it is proved or the proof cannot be continued.

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{P \to Q, R \Rightarrow R}}{P \to Q, \neg R, R \Rightarrow}(\neg l) \qquad \dfrac{\dfrac{R \Rightarrow P, Q \qquad \dfrac{}{Q, R \Rightarrow Q}}{P \to Q, R \Rightarrow Q}(\to l)}{P \to Q, \neg Q, R \Rightarrow}(\neg l)}{P \to Q, \neg R \lor \neg Q, R \Rightarrow}(\lor l)}{P \to Q, \neg R \lor \neg Q \Rightarrow \neg R}(\neg r)}{(P \to Q) \land (\neg R \lor \neg Q) \Rightarrow \neg R}(\land l)}{\Rightarrow [(P \to Q) \land (\neg R \lor \neg Q)] \to \neg R}(\to r)
$$

(c) Design a method for determining whether a prepositional formula is inconsistent. The method should work by examining the formula's disjunctive normal form. Demonstrate your method by applying it to the formula:

$$\neg\left[(P \land Q) \lor (Q \to P)\right]$$

Informally:

Use DNF but not canonical DNF.

- Convert the formula to DNF
- If the reduction to DNF yields **f** then the formula is inconsistent and unsatisfiable.
- If the reduction yields **t** then the formula is valid
- If the DNF is a disjunction of multiple terms, then the formula is satisfiable but not valid.

Formally:

- Start with a propositional formula F.
- Replace all occurrences of $A \leftrightarrow B$ with $(A \to B) \wedge (B \to A)$:

$$A \leftrightarrow B \Rightarrow (A \to B) \wedge (B \to A)$$

- Replace all occurrences of $A \leftarrow B$ with $B \to A$

$$A \leftarrow B \Rightarrow B \to A$$

- Replace all occurrences of $A \to B$ with $\neg A \vee B$:

$$A \to B \Rightarrow \neg A \vee B$$

- While there are any negated formulae (ie negations which are not literals); push in a negation using De Morgens Laws. IE while it can be done; , do one one of the following operations.

$$\neg(A \vee B) \Rightarrow \neg A \wedge \neg B$$
$$\neg(A \wedge B) \Rightarrow \neg A \vee \neg B$$

  The formula is now in negation normal form
- While there is a double negation, eliminate a double-negation:

$$\neg\neg A \Rightarrow A$$

- While it can be done; apply distributivity to push all $\wedge$s in.

$$A \wedge (B \vee C) \Rightarrow A \wedge B \vee A \wedge C$$

- Use associativity to remove all brackets
- Remove all conjuncts of the form $A \wedge \cdots \wedge \neg A$

No, this isn't enough to canonicalise.

OR conjuncts of the form (... & f) or conjuncts of the form (... & ¬t)

After performing these operations, the formula is in DNF. If the formula is non-empty then there exists one conjunct which was not removed – this conjunct is satisfiable. Since Skolemization preserves consistency, we can conclude that the original formula was consistent in this case.

3 2005 Paper 5 Question 9

https://www.cl.cam.ac.uk/teaching/exams/pastpapers/y2005p5q9.pdf

(a) In order to prove the following formula by resolution, what set of clauses should be submitted to the prover? Justify your answer briefly.

$$\forall x\,[P(x) \vee Q \to \neg R(x)] \wedge \forall x\,[(Q \to \neg S(x)) \to (P(x) \wedge R(x))] \to \forall x S(x)$$

We should Skolemize the formula, then convert to DNF and negate (so the negation is in CNF). The clauses we submit to the prover should be the disjunctions.

In this particular case, this formula uses implication. Therefore, we can simplify it quicker using the identity $\neg(A \to B) \simeq A \wedge \neg B$.

$$\forall x\,[P(x) \vee Q \to \neg R(x)] \wedge \forall x\,[(Q \to \neg S(x)) \to (P(x \wedge R(x)))] \to \forall x S(x)$$

Skolemize the formula. Since there are no existential quantifiers, this involves only dropping the universal quantifiers

$$[P(x) \vee Q \to \neg R(x)] \wedge [(Q \to \neg S(x)) \to (P(x) \wedge R(x))] \to S(x)$$

resolution works on CNF (clauses are disjunctions of variables and their negations)

Next negate. I will use the identity $\neg(A \to B) \simeq A \wedge \neg B$

$$
\begin{aligned}
&\simeq [P(x) \vee Q \to \neg R(x)] \wedge [(Q \to \neg S(x)) \to (P(x) \wedge R(x))] \wedge \neg S(x)\\
&\simeq [\neg(P(x) \vee Q) \vee \neg R(x)] \wedge [\neg (Q \to \neg S(x)) \wedge (P(x) \wedge R(x))] \wedge \neg S(x)\\
&\simeq [\neg P(x) \wedge Q \vee \neg R(x)] \wedge [\neg (\neg Q \vee \neg S(x)) \wedge P(x) \wedge R(x)] \wedge \neg S(x)\\
&\simeq [\neg(P(x) \vee Q) \vee \neg R(x)] \wedge [\neg (Q \to \neg S(x)) \wedge (P(x) \wedge R(x))] \wedge \neg S(x)\\
&\simeq [\neg P(x) \wedge \neg Q \vee \neg R(x)] \wedge [Q \wedge S(x) \wedge P(x) \wedge R(x)] \wedge \neg S(x)\\
&\simeq (\neg P(x) \vee \neg R(s)) \wedge (\neg Q \vee \neg R(x)) \wedge Q \wedge S(x) \wedge P(x) \wedge R(x) \wedge \neg S(x)\\
&\simeq \{\neg P(x), \neg R(s)\}\ \{\neg Q, \neg R(x)\}\ \{Q\}\ \{S(x)\}\ \{P(x)\}\ \{R(x)\}\ \{\neg S(x)\}
\end{aligned}
$$

negate first.
skolemise
convert to CNF
read off clauses

(b) Derive the empty clause using resolution with the following set of clauses, or give convincing reasons why it cannot be derived.

$$\{\neg P(x,x)\}\ \{P(x,f(x))\}\ \{\neg P(x,y), \neg P(y,z), P(x,z)\}$$

Note that variables' scope ~~is~~ are *only* in their clauses. For clarity, I will rename the variables

$$=\{\neg P(v,v)\}\ \{P(w,f(w))\}\ \{\neg P(x,y), \neg P(y,z), P(x,z)\}$$

To apply resolution, we must unify $\neg P(v,v)$ and $P(x,z)$. Using the most general unifier $\sigma = [v/x, v/z]$

one option but not true that we /must/ do this.

$$
\begin{aligned}
=&\{\neg P(v,v)\}\ \{P(w,f(w))\}\ \{\neg P(v,y), \neg P(y,v), P(v,v)\}\\
=&\{P(w,f(w))\}\ \{\neg P(v,y), \neg P(y,v)\}
\end{aligned}
$$

Notice the two terms in the RHS clause are the same under α-equivalence. Using idempotence, we can remove one.

It's impossible. The clauses are trivially satisfied by any anti-reflexive, transitive relation where 'f' as a generator, e.g. P = < on Integers and f(x)=x+1.

$$=\{P(w,f(w))\}\ \{\neg P(y,v)\}$$

Use the unifier $\sigma = [w/y, f(x)/v]$

$$
\begin{aligned}
=&\{P(w,f(w))\}\ \{\neg P(w,f(w))\}\\
=&\square
\end{aligned}
$$

(c) Derive the empty clause using resolution with the following set of clauses or give convincing reasons why it cannot be derived. (Note that a and b are constants.)

$$\{\neg P(a)\}\ \{Q(a)\}\ \{R(b)\}\ \{S(b)\}\ \{\neg Q(x), P(x), \neg R(y), \neg Q(y)\}\ \{\neg S(x), \neg R(x), Q(x)\}$$

Since variables are scoped only within their clause, I rename for clarity

$$=\{\neg P(a)\}\ \{Q(a)\}\ \{R(b)\}\ \{S(b)\}\ \{\neg Q(x), P(x), \neg R(y), \neg Q(y)\}\ \{\neg S(z), \neg R(z), Q(z)\}$$

Using the unifier $\sigma = [a/x]$

$$
\begin{aligned}
=&\{R(b)\}\ \{S(b)\}\ \{\neg P(a)\}\ \{Q(a)\}\ \{\neg Q(a), P(a), \neg R(y), \neg Q(y)\}\ \{\neg S(z), \neg R(z), Q(z)\}\\
=&\{R(b)\}\ \{S(b)\}\ \{\neg P(a)\}\ \{P(a), \neg R(y), \neg Q(y)\}\ \{\neg S(z), \neg R(z), Q(z)\}\\
=&\{R(b)\}\ \{S(b)\}\ \{\neg R(y), \neg Q(y)\}\ \{\neg S(z), \neg R(z), Q(z)\}
\end{aligned}
$$

Using the unifier $\sigma = [b/y, b/z]$

$=\{R(b)\}\ \{S(b)\}\ \{\neg R(b), \neg Q(b)\}\ \{\neg S(b), \neg R(b), Q(b)\}$
$=\{R(b)\}\ \{\neg R(b), \neg Q(b)\}\ \{S(b)\}\ \{\neg S(b), \neg R(b), Q(b)\}$

Using idempotence, we can duplicate the clause $\{R(b)\}$

$=\{R(b)\}\ \{\neg R(b), \neg Q(b)\}\ \{R(b)\}\ \{S(b)\}\ \{\neg S(b), \neg R(b), Q(b)\}$
$=\{\neg Q(b)\}\ \{R(b)\}\ \{\neg R(b), Q(b)\}$
$=\{\neg Q(b)\}\ \{Q(b)\}$
$=\square$

Wow. What a journey.

Herbrand Universes are
- More general than any other universes
    - Assume f(a) is a COMPLETELY new point in the universe!
    - The Herbrand Universe does this repeatedly.
    - Its therefore able to mimic any universe but never gets stuck like any other universe.
- Syntactic
    - no inbuilt knowledge about any items
- There are finite refutations
    - if something ISN'T true there is a finite counterexample

This tells you the set of ground terms which falsify an expression is finite -- not that you can find it in finite time!

If something cannot be proved in the Herbrand Universe then it cannot be done in any universe.

Functions refer to different points -- ie f(x) \neq g(y) for ANY VALUES

When negating, separate the formula and do each conjunction / disjunction independently.
IE (complex expr) -> (compex expr 2) = A -> B
So turn into A n ¬B and find claues for A and ¬B separately

Clause methods try to find contradictions -- if there are subsets where P is true and other subsets where P is false then you try to find an area in the intersection of those areas.

Why is a variable only scoped within its clause?
Skolemization pulls exists out
So we have a forall of conjunctions
It's therefore logically equivalent to push the forall in!
Therefore variables are logically scoped!

When performing unification
We don't know what to unify
If we find the empty clause then we show that the clauses are contradictory
In non-trivial examples then if we DON'T find the empty clause -- this means nothing!

A Herbrand Model is
A Model where the domain is the Herbrand Universe
The interpretation:
all constants map to themselves! It's important that the points they represent are the points represented by their names
ie {c -> c, f -> f, g -> g, ...}
predicates can be defined by naming all points in the universe for which it holds
ie {{ = {c}, Q = {c}}

Conversion to DNF:
- push in ¬
- distribute and over or ie A (B v C) -> AB v AC

Resolution ADDS new clauses -- it does not delete clauses
it just adds new shorter clauses
which you can then use to try and form contradictions

To prove that the empty clause cannot be derived:
- provide an example that satisfies it
- perform a prolog-style exptime DFS