Logical reasoning and programming, lab session 8

(November 8, 2021)

8.1 Derive the empty clause $\Box$ using the resolution calculus from:

(a) $\{\{\neg p(X), \neg p(f(X))\}, \{p(f(X)), p(X)\}, \{\neg p(X), p(f(X))\}\}$

(b) $\{\{\neg p(X, a), \neg p(X, Y), \neg p(Y, X)\}, \{p(X, f(X)), p(X, a)\}, \{p(f(X), X), p(X, a)\}\}$

8.2 Prove using the resolution calculus that from

$$\begin{aligned}
&\forall X \forall Y (p(X, Y) \rightarrow p(Y, X)) \\
&\forall X \forall Y \forall Z ((p(X, Y) \wedge p(Y, Z)) \rightarrow p(X, Z)) \\
&\forall X \exists Y p(X, Y)
\end{aligned}$$

follows $\forall X p(X, X)$.

8.3 Check PyRes; simple resolution-based theorem provers for first-order logic. You can find proofs for the previous examples using them.

8.4 List all the possible applications of the factoring rule on the clause

$$\{p(X, f(Y), Z), p(T, T, g(a)), p(f(b), S, g(W)), \neg s(Z, T), \neg s(c, d)\}.$$

If it is possible to use the factoring rule several times, then produce even these results.

8.5 Produce all the possible paramodulants, but do not perform paramodulations into variables, of

$$\{\{p(X), \neg q(X, Y), f(c, Y) = g(X)\}, \{p(Z), q(g(a), f(Z, b)), c = f(c, c)\}\}.$$

8.6 Formulate the following problems in the TPTP language and (dis)prove them using the E prover. Assuming the following group axioms

$$\begin{aligned}
&e \cdot X = X, \\
&X^{-1} \cdot X = e, \\
&(X \cdot Y) \cdot Z = X \cdot (Y \cdot Z)
\end{aligned}$$

your task is to (dis)prove

(a) $X \cdot e = X$,

(b) $X \cdot X^{-1} = e$,

(c) $X \cdot Y = Y \cdot X$,

(d) $X \cdot Y = Y^{-1} \cdot X^{-1}$.

8.7 Use the model finder Paradox to produce counterexamples for unprovable claims in the previous exercise **8.6**.

8.8 Use PyRes to prove **8.6**a. Note that PyRes uses the naïve handling of equality. For example, use

```
pyres-fof.py -tifb -HPickGiven5 -nlargest
```

There are various heuristics (`FIFO`, `SymbolCount`, `PickGiven5`, and `PickGiven2`) and literal selections (`first`, `smallest`, `largest`, `leastvars`, and `eqleastvars`) available.

8.9 Formalize in the TPTP format a simple example with the following axioms

$$\forall X \neg r(X, X),$$
$$\forall X \forall Y \forall Z (r(X, Y) \wedge r(Y, Z) \rightarrow r(X, Z)),$$
$$\forall X \exists Y r(X, Y)$$

and check how fast can Paradox generate possible finite models for this simple problem. Clearly, it will never find a model, because the problem has only infinite models.

8.10 Try the Vampire prover on the problem GRP140-1 from the TPTP library. We demonstrate the effect of the limited resource strategy (LRS), which discards unprocessed clauses that will be unlikely processed in a given time limit, by this example. For the intended behavior you need a special setting—age:weight ratio is 5:1 and the forward subsumption is turned off:

```
vampire -awr 5:1 -fsr off -t 30 GRP140-1.p
```

First, try the timelimit 30s, then try 15s, 7s, ... You can try even shorter times than 1s, e.g., `-t 5d` means 5 deciseconds.

For comparison you can try the competition mode on the same problem

```
vampire --mode casc GRP140-1.p
```

8.11 Try the E prover on the problem GRP001-1 from the TPTP library. Compare how can the use of a literal selection strategy influence the behavior of the prover:

```
eprover --literal-selection-strategy=NoSelection GRP001-1.p
eprover --literal-selection-strategy=SelectLargestNegLit \
          GRP001-1.p
```

You may also visualize the proof using the Interactive Derivation Viewer (IDV) tool for graphical rendering of derivations through System on TPTP.