Automated Theorem Proving
(Demonstrarea Automata a Teoremelor)
Resolution

Mădălina Eraşcu[1] and Tudor Jebelean[2]

[1]West University of Timişoara and Institute e-Austria Timişoara, bvd. V. Parvan 4, Timişoara, Romania,
merascu@info.uvt.ro

[2]Research Institute for Symbolic Computation, Johannes Kepler University, Linz, Austria,
tjebelea@risc.jku.at

May 26, 2014

Example 1 Prove by resolution that G is a logical consequence of F_1 and F_2 where

$$\begin{array}{ll} F_1: & \underset{x}{\forall}(C[x] \;\Rightarrow\; (W[x] \,\wedge\, R[x])) \\ F_2: & \underset{x}{\exists}(C[x] \wedge O[x]) \\ G: & \underset{x}{\exists}(O[x] \wedge R[x]) \end{array}$$

Solution. We show that $F_1 \wedge F_2 \wedge \neg G$ is unsatisfiable by resolution. We transform F_1, F_2, $\neg G$ into Skolem standard form. We have

$$\begin{array}{rl} F_1: & \underset{x}{\forall}(C[x] \;\Rightarrow\; (W[x] \,\wedge\, R[x])) \\ \Longleftrightarrow & \underset{x}{\forall}(\neg C[x] \;\vee\; (W[x] \,\wedge\, R[x])) \\ \Longleftrightarrow & \underset{x}{\forall}(\neg C[x] \,\vee\, W[x]) \,\wedge\, (\neg C[x] \,\vee\, R[x]) \end{array}$$

$$\begin{array}{rl} F_2: & \underset{x}{\exists}(C[x] \wedge O[x]) \\ \rightsquigarrow & C[a] \wedge O[a] \end{array}$$

$$\begin{array}{rl} \neg G: & \neg\left(\underset{x}{\exists}(O[x] \wedge R[x])\right) \\ \Longleftrightarrow & \underset{x}{\forall}(\neg O[x] \vee \neg R[x]) \end{array}$$

We have the following set of clauses

$$\begin{array}{ll} (1) & \neg C[x] \,\vee\, W[x] \\ (2) & \neg C[x] \,\vee\, R[x] \\ (3) & C[a] \\ (4) & O[a] \\ (5) & \neg O[x] \vee \neg R[x] \end{array}$$

By resolution we obtain also the following clauses

$$\begin{array}{lll} (6) & \neg R[a] & (4) \wedge (5), \{x \to a\} \\ (7) & \neg C[a] & (6) \wedge (2), \{x \to a\} \\ (8) & \emptyset & (7) \wedge (3) \end{array}$$

◄

Example 2 Prove by resolution that G is a logical consequence of F_1 and F_2 where

$$\begin{array}{ll} F_1: & \underset{x}{\exists}\left(P[x] \;\wedge\; \underset{y}{\forall}(D[y] \;\Rightarrow\; L[x,y])\right) \\ F_2: & \underset{x}{\forall}\left(P[x] \;\Rightarrow\; \underset{y}{\forall}(Q[y] \;\Rightarrow\; \neg L[x,y])\right) \\ G: & \underset{x}{\forall}(D[x] \;\Rightarrow\; \neg Q[x]) \end{array}$$

Solution. We show that $F_1 \wedge F_2 \wedge \neg G$ is unsatisfiable by resolution. We transform F_1, F_2, $\neg G$ into Skolem standard form. We have

$$\begin{array}{rl} F_1: & \underset{x}{\exists}\left(P[x] \;\wedge\; \underset{y}{\forall}(D[y] \;\Rightarrow\; L[x,y])\right) \\ \Longleftrightarrow & \underset{x}{\exists}\left(P[x] \;\wedge\; \underset{y}{\forall}(\neg D[y] \;\vee\; L[x,y])\right) \\ \Longleftrightarrow & \underset{x\,y}{\exists\forall}(P[x] \;\wedge\; (\neg D[y] \;\vee\; L[x,y])) \\ \rightsquigarrow & \underset{y}{\forall}(P[a] \;\wedge\; (\neg D[y] \;\vee\; L[a,y])) \end{array}$$

$$\begin{array}{rl} F_2: & \underset{x}{\forall}\left(P[x] \;\Rightarrow\; \underset{y}{\forall}(Q[y] \;\Rightarrow\; \neg L[x,y])\right) \\ \Longleftrightarrow & \underset{x}{\forall}\left(P[x] \;\Rightarrow\; \underset{y}{\forall}(\neg Q[y] \;\vee\; \neg L[x,y])\right) \\ \Longleftrightarrow & \underset{x}{\forall}\left(\neg P[x] \;\vee\; \underset{y}{\forall}(\neg Q[y] \;\vee\; \neg L[x,y])\right) \\ \Longleftrightarrow & \underset{x\,y}{\forall\forall}(\neg P[x] \;\vee\; \neg Q[y] \;\vee\; \neg L[x,y]) \end{array}$$

$$\begin{array}{rl} \neg G: & \neg\left(\underset{x}{\forall}(D[x] \;\Rightarrow\; \neg Q[x])\right) \\ \Longleftrightarrow & \neg\left(\underset{x}{\forall}(\neg D[x] \;\vee\; \neg Q[x])\right) \\ \Longleftrightarrow & \underset{x}{\exists}(D[x] \;\wedge\; Q[x]) \\ \rightsquigarrow & D[a] \;\wedge\; Q[a] \end{array}$$

We have the following set of clauses

$$\begin{array}{ll} (1) & P[a] \\ (2) & \neg D[y] \;\vee\; L[a,y] \\ (3) & \neg P[x] \;\vee\; \neg Q[y] \;\vee\; \neg L[x,y] \\ (4) & D[a] \\ (5) & Q[a] \end{array}$$

By resolution we obtain also the following clauses

$$\begin{array}{lll}
(6) & L[a,a] & (2) \wedge (4), \{y \to a\} \\
(7) & \neg P[a] \ \vee \ \neg Q[a] & (3) \wedge (6), \{x \to a, y \to a\} \\
(8) & \neg Q[a] & (1) \wedge (7) \\
(9) & \emptyset & (5) \wedge (8)
\end{array}$$

◀

Example 3 Prove by resolution that G is a logical consequence of F where

$$\begin{array}{ll}
F: & \underset{x}{\forall}\underset{y}{\exists}(S[x,y] \ \wedge \ M[y]) \quad \Rightarrow \quad \underset{y}{\exists}(I[y] \ \wedge \ E[x,y]) \\
G: & \neg\underset{x}{\exists} I[x] \quad \Rightarrow \quad \underset{x,y}{\forall}\ (S[x,y] \Rightarrow \neg M[y])
\end{array}$$

Solution. We show that $F \wedge \neg G$ is unsatisfiable. First we transform the formulas into standard form. We have

$$\begin{array}{ll}
F: & \underset{x}{\forall}\left(\underset{y}{\exists}(S[x,y] \ \wedge \ M[y])\right) \quad \Rightarrow \quad \underset{y}{\exists}(I[y] \ \wedge \ E[x,y]) \\
\Longleftrightarrow & \underset{x}{\forall}\neg\left(\underset{y}{\exists}(S[x,y] \ \wedge \ M[y])\right) \ \vee \ \underset{y}{\exists}(I[y] \ \wedge \ E[x,y]) \\
\Longleftrightarrow & \underset{x}{\forall}\left(\underset{y}{\forall}(\neg S[x,y] \ \vee \ \neg M[y])\right) \ \vee \ \underset{y}{\exists}(I[y] \ \wedge \ E[x,y]) \\
\Longleftrightarrow & \underset{x}{\forall}\left(\underset{y}{\forall}(\neg S[x,y] \ \vee \ \neg M[y])\right) \ \vee \ (I[f[x]] \ \wedge \ E[x,f[x]]) \\
\Longleftrightarrow & \underset{x}{\forall}\underset{y}{\forall}(\neg S[x,y] \ \vee \ \neg M[y]) \ \vee \ (I[f[x]] \ \wedge \ E[x,f[x]]) \\
\Longleftrightarrow & \underset{x}{\forall}\underset{y}{\forall}((\neg S[x,y] \ \vee \ \neg M[y] \ \vee \ I[f[x]]) \ \wedge \ (\neg S[x,y] \ \vee \ \neg M[y] \ \vee \ E[x,f[x]])) \\
\neg G: \neg & \left(\neg\underset{x}{\exists} I[x] \quad \Rightarrow \quad \underset{x,y}{\forall}(S[x,y] \Rightarrow \neg M[y])\right) \\
\Longleftrightarrow & \neg\left(\neg\underset{x}{\exists} I[x] \quad \Rightarrow \underset{x,y}{\forall}(\neg S[x,y] \vee \neg M[y])\right) \\
\Longleftrightarrow & \neg\left(\underset{x}{\exists} I[x] \quad \vee \quad \underset{x,y}{\forall}(\neg S[x,y] \vee \neg M[y])\right) \\
\Longleftrightarrow & \left(\underset{x}{\forall}\neg I[x] \quad \wedge \quad \underset{x,y}{\exists}(S[x,y] \wedge M[y])\right) \\
\Longleftrightarrow & \underset{z}{\forall}\neg I[z] \quad \wedge \quad \underset{x,y}{\exists}(S[x,y] \wedge M[y]) \\
\rightsquigarrow & \underset{z}{\forall}\neg I[z] \quad \wedge \quad S[a,b] \wedge M[b]
\end{array}$$

We have the following set of clauses

$$\begin{array}{ll}
(1) & \neg S[x,y] \ \vee \ \neg M[y] \ \vee \ I[f[x]] \\
(2) & \neg S[x,y] \ \vee \ \neg M[y] \ \vee \ E[x,f[x]] \\
(3) & \neg I[z] \\
(4) & S[a,b] \\
(5) & M[b]
\end{array}$$

By resolution we obtain also the following clauses

$$\begin{array}{lll}
(6) & \neg S[x,y] \ \vee \ \neg M[y] & (1) \wedge (3), \{z \to f[x]\} \\
(7) & \neg M[b] & (4) \wedge (6), \{x \to a, y \to b\} \\
(8) & \emptyset & (5) \wedge (7)
\end{array}$$

◀

Example 4 Prove by resolution that G is a logical consequence of F_1, F_2, and F_3 where

$$\begin{aligned}
F_1 :&\quad \underset{x}{\forall}\,(Q[x] \;\Rightarrow\; \neg P[x]) \\
F_2 :&\quad \underset{x}{\forall}\left((R[x] \;\wedge\; \neg Q[x]) \quad\Rightarrow\quad \underset{y}{\exists}\,(T[x,y] \;\wedge\; S[y])\right) \\
F_3 :&\quad \underset{x}{\exists}\left(P[x] \;\wedge\; \underset{y}{\forall}\,(T[x,y] \;\Rightarrow\; P[y]) \;\;\wedge\;\; R[x]\right) \\
G :&\quad \underset{x}{\exists}\,(S[x] \wedge P[x])
\end{aligned}$$

Solution. We show that $F_1 \wedge F_2 \wedge F_3 \wedge \neg G$ is unsatisfiable. First we transform the formulas into standard form.

$$\begin{aligned}
F_1 :&\; \underset{x}{\forall}\,(Q[x] \;\Rightarrow\; \neg P[x]) \quad\iff\quad \underset{x}{\forall}\,(\neg Q[x] \;\vee\; \neg P[x]) \\
F_2 :&\; \underset{x}{\forall}\left((R[x] \;\wedge\; \neg Q[x]) \quad\Rightarrow\quad \underset{y}{\exists}\,(T[x,y] \;\wedge\; S[y])\right) \\
&\iff\quad \underset{x}{\forall}\left(\neg\,(R[x] \;\wedge\; \neg Q[x]) \quad\vee\quad \underset{y}{\exists}\,(T[x,y] \;\wedge\; S[y])\right) \\
&\iff\quad \underset{x}{\forall}\left(\neg R[x] \;\vee\; Q[x] \quad\vee\quad \underset{y}{\exists}\,(T[x,y] \;\wedge\; S[y])\right) \\
&\iff\quad \underset{x}{\forall}\underset{y}{\exists}\,(\neg R[x] \;\vee\; Q[x] \quad\vee\quad (T[x,y] \;\wedge\; S[y])) \\
&\iff\quad \underset{x}{\forall}\underset{y}{\exists}\,((\neg R[x] \;\vee\; Q[x] \quad\vee\quad T[x,y]) \;\;\wedge\;\; (\neg R[x] \;\vee\; Q[x] \;\vee\; S[y])) \\
&\rightsquigarrow\; \underset{x}{\forall}\,((\neg R[x] \;\vee\; Q[x] \quad\vee\quad T[x,f[x]]) \;\;\wedge\;\; (\neg R[x] \;\vee\; Q[x] \;\vee\; S[f[x]])) \\
F_3 :&\; \underset{x}{\exists}\left(P[x] \;\wedge\; \underset{y}{\forall}\,(T[x,y] \;\Rightarrow\; P[y]) \;\;\wedge\;\; R[x]\right) \\
&\iff\quad \underset{x}{\exists}\left(P[x] \;\wedge\; \underset{y}{\forall}\,(\neg T[x,y] \;\vee\; P[y]) \;\;\wedge\;\; R[x]\right) \\
&\iff\quad \underset{x}{\exists}\underset{y}{\forall}\,(P[x] \;\wedge\; (\neg T[x,y] \;\vee\; P[y]) \;\;\wedge\;\; R[x]) \\
&\rightsquigarrow\; \underset{y}{\forall}\,(P[a] \;\wedge\; (\neg T[a,y] \;\vee\; P[y]) \;\;\wedge\;\; R[a]) \\
\neg G :&\; \neg\left(\underset{x}{\exists}\,(S[x] \wedge P[x])\right) \\
&\iff\quad \underset{x}{\forall}\,(\neg S[x] \vee \neg P[x])
\end{aligned}$$

We have the following set of clauses

(1)	$\neg Q[x] \ \vee \ \neg P[x]$	
(2)	$\neg R[x] \ \vee \ Q[x] \ \ \vee \ \ T[x, f[x]]$	
(3)	$\neg R[x] \ \vee \ Q[x] \ \vee \ S[f[x]]$	
(4)	$P[a]$	
(5)	$\neg T[a, y] \ \vee \ P[y]$	
(6)	$R[a]$	
(7)	$\neg S[x] \vee \neg P[x]$	
(8)	$\neg Q[a]$	$(1) \wedge (4), \{x \to a\}$
(9)	$\neg R[a] \ \vee \ T[a, f[a]]$	$(8) \wedge (2), \{x \to a\}$
(10)	$\neg R[a] \ \vee \ P[f[a]]$	$(9) \wedge (5), \{y \to f[a]\}$
(11)	$P[f[a]]$	$(10) \wedge (6)$
(12)	$\neg S[f[a]]$	$(11) \wedge (7)$
(13)	$\neg R[a] \ \vee \ Q[a]$	$(12) \wedge (3)$
(14)	$Q[a]$	$(13) \wedge (6)$
(15)	$\emptyset$	$(14) \wedge (8)$

◀