ADuct HW 3&4

3. 1) $S = \{\underset{(1)}{\neg p \vee \neg q},\ \underset{(2)}{\neg p \vee q},\ \underset{(3)}{p \vee \neg q},\ \underset{(4)}{p \vee q}\}$ Show that there are infinite different BR derivations of the $\square$ clause.

For a derivation with 5 steps consider

$$\dfrac{\dfrac{\dfrac{(1)\ \neg p \vee \neg q \quad (3)\ p \vee \neg q}{(5)\ \neg q \vee \neg q}\,(BR)}{(6)\ \neg q}\,(Fact) \qquad \dfrac{\dfrac{(2)\ \neg p \vee q \quad (4)\ p \vee q}{(7)\ q \vee q}\,(BR)}{(8)\ q}\,(Fact)}{(9)\ \square}\,(BR)$$

To get a derivation with 7 steps we can replace the right part, namely $\dfrac{(2)\ \ (4)}{(7)}(BR)$ with the following $\dfrac{\dfrac{(2)\ \neg p \vee q \quad (3)\ p \vee \neg q}{(10)\ q \vee \neg q}(BR) \quad \dfrac{(2)\ \neg p \vee q \quad (4)\ p \vee q}{(7)\ q \vee q}(BR)}{(7)\ q \vee q}(BR)$. (5+2 steps because we replace one (BR) step with 3 (BR) steps.) Note that we once again have the same right part of $\dfrac{(2)\ \ (4)}{(7)}(BR)$. By replacing the new occurance of this right $\dfrac{(2)\ (4)}{(7)}(BR)$ part we get a derivation with $5+2+2=9$ steps.

Inductively we get a derivation with $5+2(n+1)$ steps by taking the one with $5+2n$ steps and replacing the top right most part (of $\dfrac{(2)\ \ (4)}{(7)}(BR)$) with the $\dfrac{\dfrac{(2)\ (3)}{(10)}(BR) \quad \dfrac{(2)\ (4)}{(7)}BR}{(7)}(BR)$ counterpart.

We have now found a derivation of the $\square$ consisting of $5+2n$ steps for all $n \in \mathbb{N}$. As they have different number of steps they are all indeed different and we have found infinite different BR derivations of the $\square$.

$\square$

ADuct HW 3&4

3.2) KBO ordering $\succ$ generated by $f \gg g \gg h \gg c \gg b \gg a$ and $w(f)=0, w(g)=w(h)=w(a)=w(b)=1$
$w(c)=3$. σ... well-behaved selection function wrt $\succ$. Show S is unsat, where

$$S=\{\underbrace{f(a)\neq b \vee f(a)\neq b \vee h(b)=c}_{(1)},\ \underbrace{f(a)=b \vee h(b)=c}_{(2)},\ \underbrace{f(g(b,a))\neq c \vee h(b)\neq c}_{(3)},\ \underbrace{f(g(b,a))=c}_{(4)}\}$$

by applying saturation on S using an inference process based on ground superposition calculus.

We start by working out the weights of the ground terms:

$|f(a)|=w(f)+|a|=0+w(a)=1 \qquad |b|=w(b)=1 \qquad |h(b)|=w(h)+|b|=1+w(b)=2$

$|c|=w(c)=3 \qquad |f(g(b,a))|=w(f)+|g(b,a)|=0+w(g)+|b|+|a|=1+w(b)+w(a)=3$

As $f \gg b$ and $f \gg c$ together with the calculated weights we get

$f(a) > b > h(b) > f(g(b,a)) > c$. Now we can calculate $>_{lit}$:

$\{f(a), b\} >^{bag} \{f(a)\} >^{bag} \{h(b), c\} >^{bag} \{f(g(b,a)), c\} \Rightarrow f(a)=b >_{lit} h(b)=c >_{lit} f(g(b,a))=c$

- Superposition left: $l=f(a)$, $r=b$, $C=(h(b)=c)$, $s[l]=f(a)$, $t=b$, $s[r]=b$, $D=(f(a)\neq b \vee h(b)=c)$
  (i) $l>r$ (ii) $s[l]>t$ (iii) $l=r > h(b)=c$

$$\frac{(2)\ \underline{f(a)=b} \vee h(b)=c \qquad (1)\ \underline{f(a)\neq b} \vee f(a)\neq b \vee h(b)=c}{(5)\ \underline{b\neq b} \vee h(b)=c \vee f(a)\neq b \vee h(b)=c}\ (Sup\text{-}\ell)$$

- Equality Resolution: $s=b$ $C=(h(b)=c \vee f(a)\neq b \vee h(b)=c)$

$$\frac{(5)\ \underline{b\neq b} \vee h(b)=c \vee f(a)\neq b \vee h(b)=c}{(6)\ h(b)=c \vee \underline{f(a)\neq b} \vee h(b)=c}\ (ER)$$

- Superposition left: $l=f(g(b,a))$, $r=c$, $s[l]=f(g(b,a))$, $s(r)=c$, $t=c$, $C=\square$, $D=(h(b)\neq c)$
  (i) $l>r$ (ii) $s(l)>t$ (iii) $l=r>\square$

$$\frac{(4)\ \underline{f(g(b,a))=c} \qquad (3)\ \underline{f(g(b,a))\neq c} \vee h(b)\neq c}{(7)\ \underline{c\neq c} \vee h(b)\neq c}\ (Sup\text{-}\ell)$$

- Equality Resolution: $s=c$ $C=(h(b)\neq c)$

$$\frac{(7)\ \underline{c\neq c} \vee \underline{h(b)\neq c}}{(8)\ \underline{h(b)\neq c}}\ (ER)$$

- Superposition left: $l=f(a)$, $r=b$, $s(l)=f(a)$, $t=b$, $s(r)=b$, $C=(h(b)=c)$, $D=(h(b)=c \vee h(b)=c)$
  (i) $l>r$ (ii) $s(l)>t$ (iii) $l=r> h(b)=c \in C$

$$\frac{(2)\ \underline{f(a)=b} \vee h(b)=c \qquad (6)\ h(b)=c \vee \underline{f(a)\neq b} \vee h(b)=c}{(9)\ \underline{b\neq b} \vee h(b)=c \vee h(b)=c \vee h(b)=c}\ (Sup\text{-}\ell)$$

...

ADuct HW 3&4

- 3.2.) ... Equality Resolution: $s=b$, $C = h(b)=c \lor h(b)=c \lor h(b)=c$

$$\frac{(9)\ \underline{b \neq b} \lor h(b)=c \lor h(b)=c \lor h(b)=c}{(10)\ \underline{h(b)=c} \lor \underline{h(b)=c} \lor \underline{h(b)=c}}\ (ER)$$

- Equality Factoring: $s=h(b)$, $t=c$, $t'=c$, $C=(h(b)=c)$ $\quad$ (i) $s>t\geq t'$, $(s=t)\geq(h(b)=c)$.

$$\frac{(10)\ \underline{h(b)=c} \lor \underline{h(b)=c} \lor h(b)=c}{(11)\ h(b)=c \lor \underline{c\neq c} \lor h(b)=c}\ (EF)$$

- Equality Resolution: $s=c$, $C=(h(b)=c \lor h(b)=c)$

$$\frac{(11)\ h(b)=c \lor \underline{c\neq c} \lor h(b)=c}{(12)\ \underline{h(b)=c} \lor \underline{h(b)=c}}\ (ER)$$

- Equality Factoring: $s=h(b)$, $t=c$, $t'=c$, $C=\square$ $\quad$ (i) $s>t\geq t'$, $(s=t)\geq \square$

$$\frac{(12)\ \underline{h(b)=c} \lor \underline{h(b)=c}}{(13)\ h(b)=c \lor \underline{c\neq c}}\ (EF)$$

- Equality Resolution: $s=c$, $C=(h(b)=c)$

$$\frac{(13)\ h(b)=c \lor \underline{c\neq c}}{(14)\ \underline{h(b)=c}}\ (ER)$$

- Superposition left: $l=h(b)$, $r=c$, $s(l)=h(b)$, $s(r)=c$, $t=c$, $C=\square$, $D=\square$

(i) $l>r$ $\quad$ (ii) $s(l)>t$ $\quad$ (iii) $s(l)=t>\square$

$$\frac{(14)\ \underline{h(b)=c} \qquad (8)\ \underline{h(b)\neq c}}{(15)\ \underline{c\neq c}}\ (Sup\text{-}left)$$

- Equality Resolution: $s=c$, $C=\square$

$$\frac{(15)\ \underline{c\neq c}}{(16)\ \underline{\square}}\ (ER)$$

As we have derived the false clause $\square$, we now know that S is unsat.

The selected literals are always underlined above, according to the ordering we have found in the first part of this exercise.

Maximal terms are always named l and $s(l)$.

ADuct HW 3 & 4

3.3.) $S = \{ f(x) = g(a),\ g(y) \neq f(b) \}$

a) Give a superposition refutation of S.

$$\frac{\overbrace{g(a)}^{l} = \overbrace{f(x)}^{r} \qquad \overbrace{g(y)}^{s[l']} \neq \overbrace{f(b)}^{t}}{\underbrace{f(x) \neq f(b)}_{(s[r] \neq t \vee C \vee D)\theta}}\ (\text{Sup-left})$$

We check the conditions:
1. $\theta = \{y \mapsto a\}$ is a mgu of $l = g(a)$ and $l' = g(y)$
2. $l' = g(y)$ is not a variable
3. $r\theta = f(x) \not\succ l\theta = g(a)$, but we ignore ordering
4. $t\theta = f(b) \not\succ s[l']\theta \doteq g(a)$, but we ignore ordering

$\Rightarrow S' = \{ f(x) = g(a),\ g(a) \neq f(b),\ f(x) \neq f(b) \}$

$$\frac{f(x) \neq f(b)}{\square}\ (\text{ER})$$

We have $s = f(x)$, $s' = f(b)$, $C = \square$, $\theta = \{x \mapsto b\}$ is a mgu of $f(x) = s$ and $s' = f(b)$.

We have reached $\square$ and thus have found a refutation.

---

b) Give a superposition refutation of S such that, in every step of your proof, the premises and the conclusion of that inference step do not use the symbol a together with b.

For clarity we asign a number to every clause: ① $f(x) = g(a)$ ② $g(y) \neq f(b)$

we rename the variable as it is different to x of the left ① clause

$$\frac{① f(x) = g(a) \qquad ① f(z) = g(a)}{③ f(x) = f(z)}\ (\text{Sup-Right});\ \theta = \{\};\ \text{only uses a and not b}$$

$$\frac{③ f(x) = f(z) \qquad ② g(y) \neq f(b)}{④ f(z) \neq g(y)}\ (\text{Sup-Left});\ \theta = \{x \mapsto b\};\ \text{only uses b}$$

$$\frac{① f(x) = g(a) \qquad ④ f(z) \neq g(y)}{⑤ g(a) \neq g(y)}\ (\text{Sup-left});\ \theta = \{x \mapsto z\};\ \text{only uses a}$$

$$\frac{⑤ g(a) \neq g(y)}{⑥ \square}\ (\text{ER})\quad \theta = \{y \mapsto a\}\ \text{only uses a}$$

$\square$

ADuct HW 3&4

3.4.) Apply unification algorithm on following formulas, also give mgu if exists.

a) $p(g(y,x), f(x,b))$ $p(z, f(g(a,b), z))$

- $p(g(y,x), f(x,b)) = p(z, f(g(a,b), z))$ is not isolated, but is of the last case ($f(s_1, \dots, s_n), f(t_1, \dots, t_n)$). Therefore we replace the equation by:

- $\{g(y,x) = z, f(x,b) = f(g(a,b), z)\}$. $g(y,x) = z$ is not isolated (z appears again) and is of the form (t, x) we switch the two sides to $z = g(y,x)$ and note that z does not occur in $g(y,x)$ therefore we replace z by $g(y,x)$ in all other equations giving us
- $\{z = g(y,x), f(x,b) = f(g(a,b), g(y,x))\}$. The second equation is not isolated, but of the form $f(s_1, \dots, s_n) = f(t_1, \dots, t_n)$. So we replace it by the set of equations $\{x = g(a,b), b = g(y,x)\}$ which gives us:
- $\{z = g(y,x), x = g(a,b), b = g(y,x)\}$. $x = g(a,b)$ is not isolated and of the form $x = t$. x does not occur in $g(a,b)$ so we replace x by $g(a,b)$ in all other equations.
- $\{z = g(y, g(a,b)), x = g(a,b), b = g(y, g(a,b))\}$. The last equation is not isolated and of the form $c = f(t_1, \dots, t_n)$. Therefore we halt with failure and there exists no unification.

---

b) $\{p(g(a,x), f(x,y)) = p(z, f(g(a,b), z))\}$. The equation is not isolated and we again are in the last case. By replacing it with the corresponding equations we get:
- $\{g(a,x) = z, f(x,y) = f(g(a,b), z)\}$. Again the first equation is not isolated because z appears also in the second equation. By case $t = x$ we replace z by $g(a,x)$ and get
- $\{z = g(a,x), f(x,y) = f(g(a,b), g(a,x))\}$. Now the second equation is not isolated and of the form $f(s_1, \dots, s_n) = f(t_1, \dots, t_n)$. We replace it with $x = g(a,b)$ and $y = g(a,x)$.
- $\{z = g(a,x), x = g(a,b), y = g(a,x)\}$. The second equation is not isolated because x appears elsewhere. We have to replace x by $g(a,b)$ in all other equations.
- $\{z = g(a, g(a,b)), x = g(a,b), y = g(a, g(a,b))\}$ Now all equations are isolated (x, y, z appear only once and all equations are of the form $x = t$). We return $\{z \mapsto g(a, g(a,b)), x \mapsto g(a,b), y \mapsto g(a, g(a,b))\}$ as the most general unifier.

ADuct HW 3&4

4.1) $p_1 \succ p_2 \succ p_3 \succ p_4$ Order the following clauses: $p_1 \vee \neg p_2$, $\neg p_1 \vee p_4$, $p_2 \vee \neg p_3$

We extend $\succ$ over literals: $\neg p_1 \succ p_1 \succ \neg p_2 \succ p_2 \succ \neg p_3 \succ p_3 \succ \neg p_4 \succ p_4$.

$\{p_1, \neg p_2\}$ corresponds to $p_1 \vee \neg p_2$, $\{\neg p_1, p_4\}$ to $\neg p_1 \vee p_4$ and $\{p_2, \neg p_3\}$ to $p_2 \vee \neg p_3$.

$\{\neg p_1, p_4\} \succ_{bag} \{p_1, \neg p_2, p_4\} \succ_{bag} \{p_1, \neg p_2\}$, because for the first $\succ_{bag}$ we have $\neg p_1 \succ p_1$ and $\neg p_1 \succ \neg p_2$ and for the second $\succ_{bag}$ we have $\forall y \in \emptyset : p_4 \succ y$.

Therefore $\neg p_1 \vee p_4 \succ p_1 \vee \neg p_2$.

$\{p_1, \neg p_2\} \succ_{bag} \{p_2, \neg p_3, \neg p_2\} \succ_{bag} \{p_2, \neg p_3\}$, because $p_1 \succ p_2$ and $p_1 \succ \neg p_3$ for the first $\succ_{bag}$ and $\forall y \in \emptyset : \neg p_2 \succ y$.

Therefore $p_1 \vee \neg p_2 \succ p_2 \vee \neg p_3$.

As $\succ_{bag}$ is transitive we get $\neg p_1 \vee p_4 \succ p_1 \vee \neg p_2 \succ p_2 \vee \neg p_3$. $\square$

ADuct HW3&4

4.2 $$\frac{f(a,b) \neq g(x) \vee f(x,y)=g(a) \qquad f(y,b)=g(y) \vee f(x,a)=g(x) \vee f(y,y)=g(z)}{f(a,a)=g(a)}$$

a) Prove above inference is sound inference of Sup (+BR).

$$\frac{(2)\ f(x,a)=g(x) \vee f(y,y)=g(z) \vee f(y,b)=g(y)}{(3)\ f(a,a)=g(a) \vee g(a) \neq g(a) \vee f(a,b)=g(a)} \text{ (EF) } \theta = \{x \mapsto a, y \mapsto a, z \mapsto a\}$$

$$\frac{(3)\ g(a) \neq g(a) \vee f(a,a)=g(a) \vee f(a,b)=g(a)}{(4)\ f(a,a)=g(a) \vee f(a,b)=g(a)} \text{ (ER) } \theta = \{\}$$

$$\frac{(1)\ f(a,b) \neq g(x) \vee f(x,y)=g(a) \qquad (4)\ f(a,b)=g(a) \vee f(a,a)=g(a)}{(5)\ f(a,y)=g(a) \vee f(a,a)=g(a)} \text{ (BR) } \theta = \{x \mapsto a\}$$

$$\frac{(5)\ f(a,y)=g(a) \vee f(a,a)=g(a)}{(6)\ f(a,a)=g(a)} \text{ (Fact) } \theta = \{y \mapsto a\}$$

We have found an inference only consisting of inference steps in the Sup+BR system, which is sound. Therefore $f(a,a)=g(a)$ follows from the two given formulas, and is thus a sound inference.

b) Is above a simplifying inference of Sup?

No, because $f(a,a)=g(a) \wedge (f(a,b) \neq g(x) \vee f(x,y)=g(a)) \not\Rightarrow (f(y,b)=g(y) \vee f(x,a)=g(x) \vee f(y,y)=g(z))$
so the right premesis does not become redundant and
$f(a,a)=g(a) \wedge (f(y,b)=g(y) \vee f(x,a)=g(x) \vee f(y,y)=g(z)) \not\Rightarrow (f(a,b) \neq g(x) \vee f(x,y)=g(a))$
so the left premesis does not become redundant. Thus the inference is a generating (= not simplifying) inference.

Note that I showed that above implications do not hold using Vampire. See files problem42_1.tptp, problem42_1.out, problem42_2.tptp and problem42_2.out. (satisfiable ⇒ no refutation was found ⇒ does not hold)

□

ADuct HW3&4

4.3) Consider the S-Sup (simultaneous superposition) rules

$$\frac{L=r \vee C \quad (s=t)[l'] \vee D}{((s=t)[r] \vee C \vee D)\theta} \qquad \frac{L=r \vee C \quad (s\neq t)[l'] \vee D}{((s\neq t)[r] \vee C \vee D)\theta}$$

where (1) θ is an mgu of l and l' (2) l' is not a variable (3) $r\theta \not\succeq l\theta$ (4) $t[l']\theta \not\succeq s[l']\theta$

Show that replacing Sup with S-Sup retains completeness.

Let S_{sup} be the saturation of S_0 using Sup and $S_{s\text{-}sup}$ be the saturation of S_0 using S-Sat. As we know that Sup is complete we have that if S_0 is unsat, then $\square \in S_{sup}$. If we now show that $S_{s\text{-}sup} \supseteq S_{sup}$ then S_0 ist unsat $\Rightarrow \square \in S_{sup} \subseteq S_{s\text{-}sup}$, and therefore replacing Sup with S-Sup retains completeness.

We can limit this to show that $S^r_{s\text{-}sup}$... saturation up to redundancy using S-sat $\supseteq$ S^r_{sup}... saturation up to redundancy using Sat, as then $S_{s\text{-}sup} \supseteq S_{sup}$ follows.

Let $\dfrac{L=r \vee C \quad s[l']=t \vee D}{(s[r]=t \vee C \vee D)\theta}$ be a Sup-r step in the saturation process of S_{sup}.

We need to show that the conclusion is in $S_{s\text{-}sup}$. If t does not contain l' then the Sup-r step is a S-Sup-r step and we have nothing more to show.

If t does contain l' we show that $(s[r]=t \vee C \vee D)\theta$ is redundant wrt. $L=r \vee C$ and $((s=t)[r] \vee C \vee D)\theta$:

According to Lecture 14 Page 35 we need to show $\exists D^*$... grounding instance of $(s[r]=t \vee C \vee D)\theta$ $\exists \theta'$... substitution: $D^* \succ (L=r \vee C)\theta'$, $D^* \succeq (((s=t)[r] \vee C \vee D)\theta)\theta'$ and $(L=r \vee C)\theta' \wedge (((s=t)[r] \vee C \vee D)\theta)\theta' \Rightarrow D^*$

ADuct HW3&4

4.4. a) see files problem44.tptp and problem44.out

b) Eplainations line by line:

preprocessing:

1, 3, 4 input (2 is missing as the conjecture implies the existence of a left_inverse)

5 negated conjecture

6, 7, 8 rewrite negated conjecture

9 add/use axiom of choice (Auswahlaxiom)

10 rewrite neg conjecture again

11, 13 rewrite input into cnf

14, 16 rewrite neg conjecture into cnf

proof:

17 backward demodulation on 13, 14

$$\frac{(13)\ \underbrace{\overbrace{(X_0 \cdot X_1)}^{s[l']} \cdot X_2 = X_0 \cdot (X_1 \cdot X_2)}_{} \qquad (14)\ \overset{r}{e} = \overbrace{X_4 \cdot X_3}^{l}}{(17)\ X_0 \cdot (X_1 \cdot X_2) = e \cdot X_2} \text{ (Sup-Right); } \theta = \begin{Bmatrix} X_0 \mapsto X_4, \\ X_1 \mapsto X_3 \end{Bmatrix}$$

backward demodulation, because (13) is made redundant by adding (17)

18 forward demodulation on 17, 11

$$\frac{(17)\ \overbrace{X_0 \cdot (X_1 \cdot X_2)}^{t} = \overbrace{e \cdot X_2}^{s[l']} \qquad (11)\ \overbrace{e \cdot X_0}^{l} = \overbrace{X_0}^{r}}{(18)\ X_0 \cdot (X_1 \cdot X_2) = X_2} \text{ (Sup-Right); } \theta = \{X_0 \mapsto X_2\}$$

forward demodulation, because (17) is made redundant by adding (18)

19 forward demodulation on 18, 14

$$\frac{(18)\ X_0 \cdot (X_1 \cdot X_2) = X_2 \qquad (14)\ e = X_4 \cdot X_3}{(19)\ X_0 \cdot e = X_2} \text{ (Sup-Right) } \theta = \{X_1 \mapsto X_4, X_2 \mapsto X_3\}$$

forward demodulation, because (18) is made redundant by adding (19)

21 superposition on 19, 19

rename because variables are different

$$\frac{(19)\ X_0 \cdot e = X_2 \qquad (19)\ X_3 \cdot e = X_1}{(21)\ X_1 = X_2} \text{ (Sup-Right) } \quad \theta = \{X_3 = X_0\}$$

31 superposition on 16, 21

$$\frac{(16)\ \underbrace{sK_0}^{}\overset{t}{} \neq \overbrace{sK_1}^{s[l']} \qquad (21)\ \overset{l}{X_1} = \overset{r}{X_2}}{(31)\ \underbrace{sK_0}_{t} \neq \underbrace{X_0}_{s[r]}} \text{ (Sup-Left) } \theta = \{sK_1 \mapsto X_1\}$$

40 subsumption resolution on 31, 21 $\quad \dfrac{(31)\ sK_0 \neq X_0 \qquad (21)\ X_1 = X_2}{(40)\ \square}$ (BR) $\theta = \begin{Bmatrix} X_1 \mapsto sK_0 \\ X_2 \mapsto X_0 \end{Bmatrix}$

$\square$ subsumes everything. We have found $\square$ and have therefore found a proof of the theorem!