G53KRR 2018 answer to the second assessed exercise ex2

Derive by resolution an empty clause from the clauses C1-C5 below. Note that a is a constant, f a functional symbol, and x, y, x_1, x_2, y_1, y_2 are variables.

C1 $[\neg S(x_1), T(f(x_1), x_1)]$

C2 $[\neg S(x), \neg T(y, x), \neg L(x), \neg H(y)]$

C3 $[S(a)]$

C4 $[L(a)]$

C5 $[\neg T(x_2, y_2), \neg S(y_2), H(x_2)]$

Several derivations are possible. The simplest although not the shortest one is to keep using unit clauses (clauses with only one literal):

C6 $[T(f(a), a)]$ from C1, C3, x_1/a

C7 $[\neg T(x_2, a)), H(x_2)]$ from C3, C5, y_2/a

C8 $[H(f(a))]$ from C6, C7, $x_2/f(a)$

C9 $[\neg T(y, a), \neg L(a), \neg H(y)]$ from C2, C3

C10 $[\neg T(y, a), \neg H(y)]$ from C4, C9

C11 $[\neg H(f(a))]$ from C6, C10, $y/f(a)$

C12 [ ] from C8, C11.

A shorter one but easier to get confused uses removal of duplicates/refactoring (since we have a disjunction inside the clause, we can rewrite $\alpha(x) \vee \alpha(x)$ to $\alpha(x)$, or even $\alpha(x) \vee \alpha(y)$ to $\alpha(x)$, because x and y are universally quantified and we can unify them in both directions).

C6 $[\neg S(x), \neg T(y, x), \neg L(x), \neg T(y, y_2), \neg S(y_2)]$ from C2, C5, x_2/y

C6r $[\neg S(x), \neg T(y, x), \neg L(x)]$ from C6 by refactoring ($\neg S(x)$, $\neg S(y_2)$ can be unified both ways, and the same for $\neg T(y, x)$, $\neg T(y, y_2)$)

C7 $[\neg T(y, a), \neg L(a)]$ from C3, C6r, x/a

C8 $[\neg T(y, a)]$ from C4, C7

C9 $[\neg S(a)]$ from C1, C8, $y/f(a), x_1/a$

C10 [ ]

Common mistakes were:

substituting for terms that are not variables: as in, $f(x)/y$, or a/x

applying resolution to three clauses in one go

I don't think it happened this time, but in previous years there was a mistake when two pairs of matching literals were eliminated in one go, for example $[P(x), Q(x)]$ and $[\neg P(x), \neg Q(x)]$ resolving to $[\ ]$. This is wrong because there is no contradiction between $\forall x(P(x) \vee Q(x))$ and $\forall x(\neg P(x) \vee \neg Q(x))$: for example, $P(x)$ could be even and $Q(x)$ could be odd.