Russell's Paradox as an Illustration of the Rule of Factoring

Background

In set theory, there is an axiom, known as the "comprehension" axiom[1] which intuitively states something like the following:

> If $\alpha(x)$ is a property of x, then there exists the set of all things that are α: $\{x \mid \alpha(x)\}$. i

For instance, there exists a property of x, "x is a cat" and correspondingly a set of all cats. There exists a property of x "x is a prime number" and correspondingly a set of all prime numbers. There exists a property of x "x is a prime number divisible by 12" and correspondingly a set of all prime numbers divisible by 12 — the empty set, of course, but that's OK.

More formally, we can assert that there exists a set s such that, for all x, x is an element of x if and only if $\alpha(x)$.

$$\exists_{\mathtt{s}} \forall_{\mathtt{x}} \mathtt{E}(\mathtt{x},\mathtt{s}) \Leftrightarrow \alpha(\mathtt{x})$$

In 1901 Bertrand Russell discovered that that formulation of the comprehension axiom won't work; it is inconsistent. Let α be the property "x is a set that is not an element of x". Then s is the set of all sets that are not elements of themselves. But now the question is: is s an element of itself? And neither way works. If s is an element of itself, then by definition it shouldn't be in s; and if it isn't, then by definition it should.

So if one is trying to axiomatize set theory, you have to find some alternative formulation of the comprehension axioms. We are not concerned with that.

Resolution theorem proving

Rather, what we want to do is to use resolution theorem proving to show that the set s can't exist. And, curiously, we can do that without using any axioms!

So in terms of our algorithm:
Γ, the set of axioms, is the null set.
ϕ the statement to be proved, is the assertion that there does not exist a set of all sets that are not elements of themselves.

$\phi \equiv \neg\exists_{\mathtt{s}} \forall_{\mathtt{x}}\mathtt{E}(\mathtt{x},\mathtt{s}) \Leftrightarrow \neg\mathtt{E}(\mathtt{x},\mathtt{x})$

Our goal is to prove ϕ from Γ, the empty set. So applying the resolution theorem proving procedure:

$\Delta = \Gamma \cup \{\neg\phi\}$ is a set containing one sentence, the negation of ϕ.
$\Delta \equiv \{\exists_{\mathtt{s}} \forall_{\mathtt{x}}\mathtt{E}(\mathtt{x},\mathtt{s}) \Leftrightarrow \neg\mathtt{E}(\mathtt{x},\mathtt{x})\}$

Converting Δ to CNF gives two clauses (`Sk` is a Skolem constant, representing the hypothetical set of all sets that are not elements of themselves.).

$\{$ 1. $\mathtt{E(x,Sk)} \vee \mathtt{E(x,x)}$
2. $\neg\mathtt{E(x,Sk)} \vee \neg\mathtt{E(x,x)}$ $\}$

[1] In the Zermelo-Frankel axiomatization of set theory, this is a special case of the replacement axiom.

Applying factoring to 1 with the substitution $\sigma = \{\texttt{x}\rightarrow\texttt{Sk}\}$ gives
3. `E(Sk,Sk).`

Applying factoring to 2 with the substitution $\sigma = \{\texttt{x}\rightarrow\texttt{Sk}\}$ gives
4. `¬E(Sk,Sk).`

Applying resolution to 3 and 4 gives the null clause.