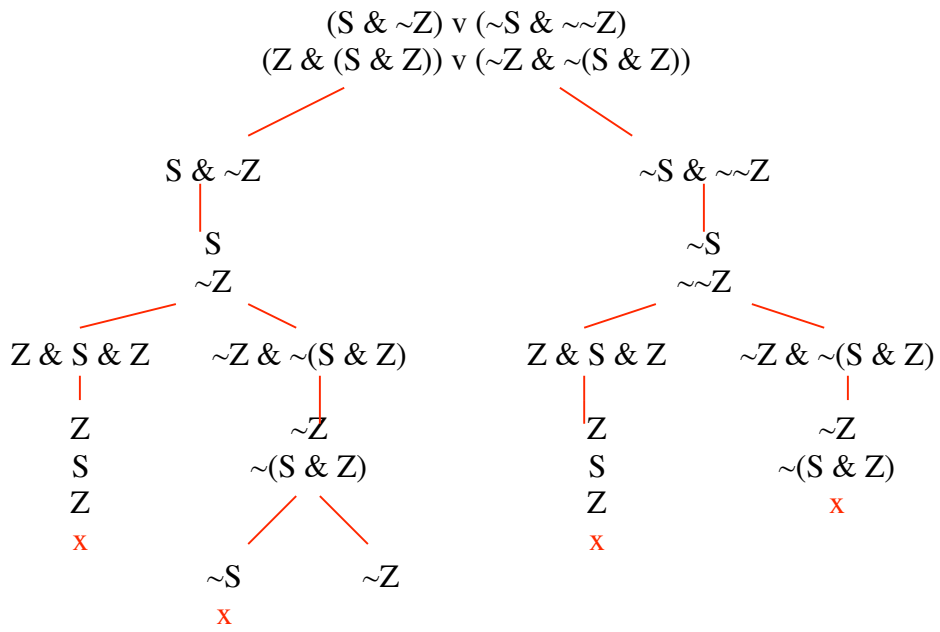
Using Semantic Tableaux to solve Knight and Knave problems

You will want to refer back to the set of rules that are also on this website for using semantic tableaux. Although that method explicitly used signs, whereas we will here use negation signs to indicate falsehood, it is pretty easy to see that the two are equivalent. When we evaluate an argument to determine its validity using this method, we write down each starting premise to indicate that it is true, and we write the negation of the conclusion to indicate that we are going to test whether the conclusion could be false. We then try to discover whether or not that was a possible starting point.

With the Knight and Knave problems we start out rather differently, since no direct argument has been given. Instead we are given a bunch of information concerning "Inhabitants of an Island" and what they say. We do not know whether the speaker is a knight or a knave (or a normal, but that is a somewhat more difficult type of puzzle). But we *do* know that *if* the speaker is a knight then s/he is speaking truthfully, and *if* the speaker is a knave then s/he is speaking falsely. If we consider just the Knight/Knave puzzles (no "normals"), then we could symbolize the claim that "A is a knight" simply by A, and the claim that "A is a knave" by ~A. And we transcribe what the speakers say into two possibilities: the speaker is a knight and what s/he says is true OR the speaker is a knave and what s/he says is false. (It is important to keep in mind what this means: if a speaker says something like "if p then q", then it is the *whole "if p then q" utterance* that is true (when the speaker is a knight) or false (when the speaker is a knave. If the speaker A says "p and q and r", then we want to represent this as: "EITHER A and (p and q and r) OR ~A and ~(p and q and r)". In some puzzles, the speaker says two (or more) *different* things - like with a period between them, or on separate occasions. Here, EACH separate sentence is true if speaker is a knight, and EACH separate sentence is false if speaker is a knave. Here is an example of a simple puzzle, and we follow this with a tableaux solution:

Sue: Zippy is a knave.
Zippy: Sue and I are knights.

Solve by a tableaux. We start with two disjunctive statements (one disjunction for each of Sue and Zippy; we use disjunctions because we do not know whether they are knights or knaves.



The branches with a red x at the bottom are impossibilities; only the open branches represent what is going on in the scenario. In this example there is only one open branch, and you can see that it contains S and ~Z. So, Sue is a knight and Zippy is a knave. Sometimes there are more than one open branch, but they contain the same information, and so that is what the answer is. Other times the two (or more) open branches contain somewhat different information, but even in this case, whatever is common to all open branches must be right. (And you just can't tell about the rest).