

COSI135: Fall 2017

Counterfactual Model Generation

In this programming assignment, you are to reason with a small subset of counterfactuals. Roughly, given a counterfactual of the form below:

If it had been the case that A, then it would have been the case that C.

this is true if and only if there is an auxiliary set S of true statements consistent with the antecedent A, such that the members of S, when conjoined with A, imply the consequent C.¹

You are to build a counterfactual model generator (CFMG) that interprets sentences such as those below in (1).

- (1) a. If John had stayed in New York, Mary would have seen him yesterday.
- b. If it hadn't rained, the show would have taken place.

There are two informational aspects to a counterfactual that need to be recovered.

1. The information that is presupposed as true. This is called the *profactual* component.
2. The information that is asserted as true. This is called the *counterfactual* component.

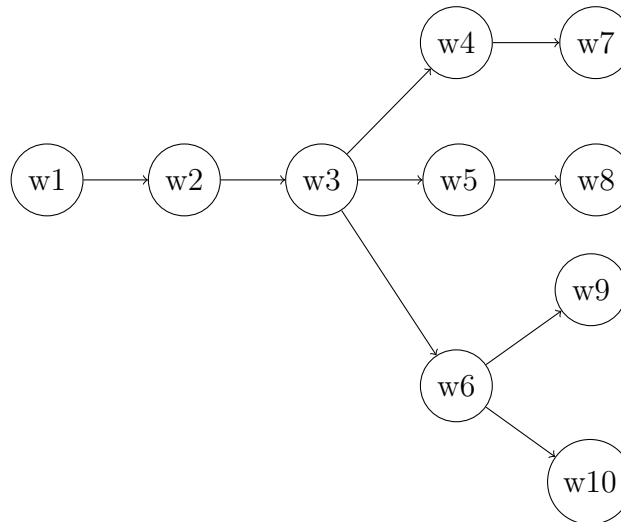
For example, for (1a) above, the profactual component is the presupposition that since John did not stay in New York, Mary did not see him.

For this assignment, you will need to write a program that performs the following three operations:

¹Menzies, Peter, "Counterfactual Theories of Causation", The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual>.

1. Modify the parser to accept the counterfactual syntax. This involves recognizing the *if*-clause as the conditional antecedent, and the following clause as its consequent. *if* is present in the lexicon of P.hs and given the category label of COND, but sentences containing “if” is not actually handled in the parser, so it is recommended that you start here.
2. Identify the profactual presupposition that accompanies the sentence, and insert the first-order expressions for this into the appropriate part of the model. The worlds must satisfy the presupposition.
3. Generate the appropriate worlds and expressions for the counterfactual component. Be sure that the worlds are partially ordered correctly and that the assertion is satisfied in the model. You will need a way to generate a model containing worlds that may branch. One way to represent branching is through a list of tuples, where each tuple is a pair representing two worlds that have an accessibility relation between them, going *from* the first member of the pair *to* the second. For example:

$[(w1, w2), (w2, w3), (w3, w4), (w3, w5), (w3, w6), (w4, w7), (w5, w8), (w6, w9), (w6, w10)]$ represents a tree structure that looks like this:



This list should be generated procedurally from the profactual and counterfactual that you parse out of your sentence. The model should be the smallest possible model that can accommodate the counterfactual.

Your submission should also contain a write-up describing how you attacked this problem, what avenues you explored, and what problems you ran into. There is no length requirement, but you should include some use cases for testing and you should describe your strategy, process, and assumptions in some depth.

NOTES:

- Part of this task is to make sure your assertion and presupposition are satisfied in the model. You should be able to modify your `isSatisfiable` or `isSatisfied` functions from PS5 to suit this part of the task (Note: your `TProp` should be satisfied in the world corresponding to the spoken time)
- The main function of this task `generateModel :: Prop -> Model` would parse a sentence, and generate the minimal model as described before. Your `Model` is an encapsulation of your `Worlds` and branching orders among them.
- This, like all the final project options, is very open ended. **It is neither required nor expected that you will develop a parser and model-generator for counterfactuals that is robust for all cases, but one that works on a few test cases will be considered very successful.** No matter how far you do get, you will be graded on the strength of your submission as it stands and on your write up, not in comparison to some ideal. This is a problem that doesn't necessarily have an established answer, and the purpose is simply to explore some of the semantic issues that arise when considering counterfactuals and their presuppositions.