Counterfactual reasoning, Pearl-style

Counterfactuals. Counterfactual reasoning involves replacing equations (and erasing arrows in the graph).

\( \text{Examples:} \)

\[
\begin{align*}
\text{cog 1 turning} & \quad \text{cog 2 turning} \\
C_2 & = 0 \\
C_3 & = 0 \\
\text{cog 1 turning} & \quad \text{cog 2 turning} \\
A & \quad B \\
C & \quad \neg C
\end{align*}
\]

This amounts to removing previous information about causal dependencies and replacing it with new information. This replacement operation is the basis of the departure from standard counterfactual logics.

Below is a model for the love triangle example. Notice that the verdicts match the judgments in the data:

\[\begin{align*}
\text{A goes} & \quad \text{B goes} \\
\text{C goes} & \quad \text{C goes}
\end{align*}\]

Implementation: Filtering Semantics

Standard counterfactual semantics (Lewis 1973, Kratzer 1981a, 1981b, 1986) we check whether all maximal consistent sets generated by adding ordering source premises to the antecedent entail the consequent.

\[\begin{align*}
[p \rightarrow q]_{\pm}^m & = 1 \text{ iff, for every maximal consistent superset } S \text{ of } (p) \text{ relative to } S', S \models [q]_{\pm}^m
\end{align*}\]

Filtering semantics (first pass): we check whether the set resulting from adding the antecedent to the ordering source and removing some relevant premise entails the consequent.

\[\begin{align*}
[p \rightarrow q]_{\pm}^m & = 1 \text{ iff } (p) \cup (q(p) \text{ filtered for } p) \text{ entails } q
\end{align*}\]

The novelty: the filtering operation, which selectively removes elements from the premise set.

The problem: this pattern cannot be vindicated by any version of premise semantics. All premise semantics validate (Kraus et al. 1990):

\[\begin{align*}
\text{LOOPT} & \quad p \quad q \\
\text{Proof for Stalnaker’s } & \quad p \quad q \\
\text{1968 semantics: the ordering relation } & \quad p \quad q \\
\text{is an ordering: it} & \quad p \quad q \\
\text{if total, hence there is a closest world to } & \quad p \quad q \\
\text{where } p & \quad q \\
\text{a-world, or an } r & \quad q
\end{align*}\]