

To determine the feasibility of a counterfactual trace, it is important to recognise two components.

First, we have to compute the probability of the sequence of event transitions. This is a difficult task, given the *Open World assumption*. In theory, we cannot know whether any event *can* follow after another event or not. However, if the data is representative of the process dynamics, we can make simplifying assumptions. For instance, we can compute the first-order transition probability by counting each transition. However, the issue remains that longer sequences tend to have a zero probability if they have never been seen in the data.

Second, we have to compute the feasibility of the individual feature values given the sequence. We can relax the computation of this probability using the *Markov Assumption*. In other words, we assume that each event vector depends on the current activity, but none of the previous events and features. Meaning, we can model density estimators for every event and use them to determine the likelihood of a set of features. **Hence, we compute the joint probability of a case by using the forward algorithm**
CITE forward algorithm .