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 events themselves. This is a difficult task, given the *open world assumption*. In theory, we cannot know whether any event *can* follow after a nother event or not. However, we can assume that the data is representative of the process dynamics. Hence, we can simply 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 . In Equation 1 shows the formulation.

$$
\begin{aligned}
feasibility_e &= p(a_i|a_{i-1} \dots a_1, \theta) \approx p(a_i|a_{i-1}, \theta) \\
feasibility_f &= p(f_i|a_i, \theta) \\
feasibility &= \prod_{i=1}^{n} [p(f_i|a_i, \theta) * p(a_i|a_{i-1}, \theta)]
\end{aligned} \tag{1}
$$

Here, a and f are the activity and features of a particular event. Likewise, θ is the data sample which is used to determine the parameters of the density function. The first equation shows the approximation based on the markov assumption.