There are many ways to estimate the density of a data set. For our purposes, we incorporate the sequential structure of the log data and make simplifying assumptions. First, we consider every activity as a state in the case. Second, each state is only dependent on its immediate predecessor and neither on future nor on any any states prior to its immediate predecessor. Third, the collection of attributes within an event depend on the activity which emits it. The second assumption is commonly known as *Markov Assumption*. With these assumptions in place, we can model the distribution by knowing the state transition probability and the density to emit a collection of event attributes given the activity.

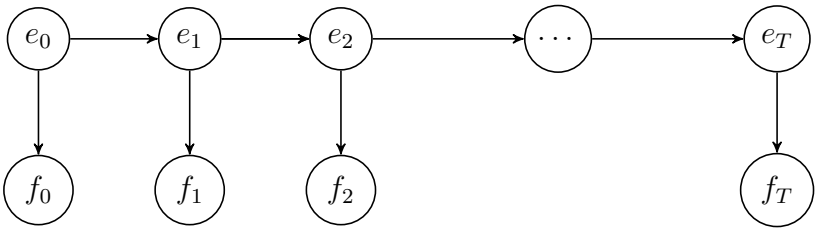

Figure 1: The feaibility model in graphical form. e_t represents an event and f_t the features it emits.

Here, e_t represents the transition from one event state to another. Likewise, f represent the emission of the feature attributes. Hence, the probability of a particular sequence is the product of the transition probability multiplied with state emission probability for each step. Note, that this is the same as the feasibility measure as in Equation 1.

$$p\left(e_{0:T}, f_{0:T}\right) = p\left(e_0\right) p\left(f_0 \mid e_0\right) \prod_{1}^{T} p\left(e_t \mid e_{t-1}\right) p\left(f_t \mid e_t\right) \tag{1}$$