Probability Over Time

by Sven Nilsen

State

A state is information associated with an object that remains true over time. We are interested in statistical knowledge about states, which can be modeled with Markov Chains. The concept explored here is how to go further beyond Markov Chains using a different approach to simulation.

Probability Over Time

When we say "probability over time" we mean the probability that an object changes state per one unit of time. It is a specific type of probability that is different from random selections etc.

Examples:

1. The probability that a married woman gives birth per year.
2. The probability that a trailer driver is involved in an accident per year.
3. The probability that a new customer calls the next 24 hours.

Motivation

We want to use computers to create experiments of probabilistic nature. Our ability to analyze such experiments depends on the theoretic underpinnings of the simulation. For example, we like to analyze the behavior of a specific system, which result is then reused in another simulation. The kind of information we are looking for is something fundamental for a large class of behavior and for different levels of complexity. The choice of “probability over time” as the type of information we look for is based on the following insights:

- Behavior tends to be described as states preserved over time, such as “walking”, “biking” etc. We like to describe the probability of leaving the state, making the default case that objects keep their behavior until the event of change occurs.
- Time is a universal dimension that can be simulated without concern of space, such that the behavior of objects in many practical applications does not need to have a physical location to be predicted.
- In many applications, the exact moment of change, sometimes the order of events, is crucial to the updated probabilities and/or interesting for the analyze. It is not sufficient to update a discrete time where multiple things happen.
- Complex systems can be modeled as a dynamical relationship of probabilities, something that excludes traditional Markov Chains as a good approximation. We also need a method to process observational data to be suitable for simulation.

We can only use probability over time to predict processes that are cyclic in behavior. From computing probability over time for simple building blocks, we can create complex processes that appears to have cycles without being strictly repetitive. With lots of simulated experiments we can study and compress knowledge gained about complex behavior.

Basic Mathematical Formulation

The probability for an object to leave state per time unit is:

$$P$$

The probability for the object to remain in same state per time unit is:

$$1-P$$

The probability for the object to remain in same state in T time units is:

$$(1-P)^T$$

The probability R for the object to leave the state after T time units is:

$$R=1-(1-P)^T$$

Solving this for T we get:

$$\boxed{T=\frac{\ln(1-R)}{\ln(1-P)}}$$

If R is same as P, the object leaves after 1 time unit. This is an important characteristics. Notice that we do not strictly define when the object is leaving, only that there is an associated probablity R of leaving after T time units.

Simulating Events

We have the equation that gives relationship between probability of leaving state per time unit (P), the time in time units (T) and the probability of leaving after that time (R).

$$T=\frac{\ln(1-R)}{\ln(1-P)}$$

If we know T or P, R can be selected randomly as a value between 0 and 1 to simulate events.

If we wait T time units and the object remains in same state, then there is an equal probability to wait T more time units. For example, the chance to get 5 coin flips ending up on head during a minute does not affect the probability of getting 5 heads the next minute. When the probability for the event is dependent on the state of the environment and the environment does not change, the probability remains the same over time. The moment we stop flipping a coin changes the environment and the probability of getting 5 heads the next minute drops to zero.

When we simulate, we can have a large complex network of objects and dependencies and pick a random number R for each of them. The smallest R gives the object that changes first, rendering all other Rs depending on that object to remain in same state invalid. When one object changes state, we update all the dependencies and pick a new R for each of them. We can pick R as many times as we like, but only if choosing R does not depend on the previous value.

Average Time

Since we know how to simulate events we can gather numerical evidence of other relations. One connection is computing the probability P of leaving state by one time unit from the average time before leaving state:

$$P = 1 - \exp\left(Avg(T)^{-1}\right)$$

From this we can derive:

$$Avg(T) = -\ln(1-P)^{-1}$$

Looking at the formula for simulating events, we find a value to put in for R that satisfy the above equation, which gives us the following connection:

$$R = 1 - \exp(-1)$$

When R is 0.5, half of the objects have changed state, but some object can remain in the same state much longer, so when we measure the average time, those object remaining in the same state for a long time moves R toward 1. When R is 1, it leads to remaining in the same state forever unless the environment changes the probability for that object to change state.

Because of such subtle differences, it is important to know whether one is looking at the problem from the P, R or T perspective.

Changing Probability Over Time

Probabilities over time are constant by nature, so when we simulate a process where the probability changes as a function of time, we need to split the time into small slices and simulate each slice.

As always when the probability depends on the relationship between objects, we can pick the minimum R for any object and update the relationship as a consequence of changing state.

Assume that we measure the frequency of events in some real world process and want to create a simulation. We set up the experiment trying to recreate the same environment as accurate as possible over and over. One object is tracked through the experiment and we start the clock the moment the object enters the state we would like to measure. When the object leaves that state, we stop the clock and write down the time. Doing this over and over in the same environment, we get a distribution of events when the object leaves the state. We split the time into slices and give each slice a number 'i'. Then we can use the following formula to compute the probability of changing state for that slice:

Algorithm 1:

$$P_i = \frac{F_i}{\sum_{j=i}^{n} F_j}$$

Computing the probability per time slice this way, it can vary up and down. When simulating we pick a random value of R. If the computed T exceed the time of the slice, we jump to the next slice and repeat the process until the object changes state.

Algorithm 2:

$$(1-P_i)^{T_i} = \frac{\sum_{j=0}^{i} F_j}{\sum F}$$

Computing the probability this way, we get a list of values that increases from 0 to 1. This list contains the probability of the object remaining in same state up to that time slice. Pick a random value of R betweeen 0 and 1 and then use binary search on the list. The index returned gives the time slice 'i' which the object changes state unless some other relation changes the probability.

Algorithm Comparison

	Simulating Event	Simulate Time Slice	Minimum Knowledge
$P_i = \frac{F_i}{\sum_{j=i}^{n} F_j}$	$O(N)$	$O(1)$	Current + Total future
$(1-P_i)^{T_i} = \frac{\sum_{j=0}^{i} F_j}{\sum F}$	$O(\log(N))$	$O(\log(N))$	Current history + Total

More About Simulation

When we have many objects and we pick a random value R for each of them, we can update the probabilities from the first event and pick new Rs, even when we ignore most of them.

Normally we can safely ignore random values as long as we do not use the information. However, in this case we pick the closest event in time. This means the information is used, but not much, so what counter-acts this effect?

The counter-acting effect is the fact that there is less probability for an object to be ignored multiple times.

TODO: Explain why using information should increase probability?

When we simulate time of change for a time slice, if the random value gives us a time beyond the slice, it does not matter which value it has, only that it is greater than the time slice. We can vary the time of the time slice. In the case of multiple objects, it happens to be the time slice equals the time to the first event.

The problem is to prove correction. A subset of a complex simulation that does not change probabilities on what happens in the rest of the system should behave the same way if simulated on its own.