# Probabilistic Agent Based Modelling: Mathematical foundations

Daniel Tang
*Deselby research institute*

## Abstract

In order to do modelling well we need to do calibration, data assimilation and uncertainty quantification properly. All of these problems can be posed in terms of probabilistic inference. At present, tools to do probabilistic inference on ABMs are lacking. In this paper I introduce the idea of a agent based dynamical system and develop a mathematical foundation for thinking about and working with probabilistic inference in these systems.

## 1 Agent based dynamical systems

We introduce the concept of an *agent based dynamical system* which consists of a set of interacting *agents*. Each agent has a state and can interact with other agents. At any given time, $t$, each agent has a propensity, $r$, to perform an action such that, in the infinitesimal slice of time, $[t,t+dt]$, there is a probability $rdt$ that the agent performs the action. The only actions an agent can perform are: change its own internal state, change the state of another agent, create a new agent or leave the system completely.

We would like to have a convenient way of specifying agent behaviours and calculating how those behaviours relate to the aggregate behaviour of the whole system. In order to do probabilistic inference we will require a way of working with probability distributions over the system (i.e. the joint state of all agents) and calculating the posterior distribution given observations of the system.

### 1.1 Single agent

We start with the simplest possible agent based dynamical system: a system with only one stateless agent. Even a system as simple as this will allow us to introduce a few key concepts. Suppose that at time $t = 0$ the agent definitely exists and between time $t$ and $t+dt$ the agent leaves the system with probability $rdt$. We now introduce a variable, $z$, to represent the state that the agent is in the system and represent the probabilistic state of the system as a polynomial in $z$

$$S = c + pz$$

where $p$ is the probability that $z$ is the case (i.e. that the agent is in the system) and $c = (1-p)$ is the probability that the system is empty. So, at time $t = 0$, $S(0) = z$. Now, we introduce an important concept, borrowed from quantum field theory, called the *annihilation* operator which acts on probabilistic states. In our case, we can definte the annihilation operator as

$$a = \frac{\partial}{\partial z}$$

If we apply this operator to $S(0)$ we get $aS(0) = az = \frac{\partial z}{\partial z} = 1$. So the annihilation operator transforms a state in which there is definitely one agent into a state which is definitely empty, i.e. the agent is annihilated. We'll see later how this generalises very nicely to the case of multiple agents.

Similarly, we define a *creation operator*:

$$a^\dagger = z$$

If we apply the creation operator to a state which is definitely empty, $S = 1$, we get $a^\dagger S = z$ which represents a state which definitely has one agent.

Given the creation and annihilation operators, it will be useful to know the rate of change of a probabilistic state if an agent has a propensity, $r$, to create or annihilate itself. Given an empty state and an agent that is currently outside of the system that has a propensity, $r$, to enter the system, then in time slice $[t,t+dt]$ there is a probability $rdt$ that the state will switch from empty to having an

agent in it. So in time $dt$ the probability of the empty state goes down by $rdt$ and the probability of there being an agent goes up by $dt$. So the rate of change of the state is givcen by

$$H_{a^\dagger} = \frac{\partial}{dt} = (a^\dagger - 1)r$$

Similarly, for the annihilation operator. If we start with an agent in the system, the probability of there being an agent decreases by $rdt$ while the probability of the state being empty increases by the same amount. We can achieve this with the operator

$$H_a = \frac{\partial}{\partial t} = (a - a^\dagger a)r$$

We'll call $H_{a^\dagger}$ and $H_a$ the *Hamiltonians* for their respective operators. Again, we'll see later that by choosing this form for the Hamiltonians we make things very elegant when we come to generalise to many agents.

Once we have the Hamiltonian for a system, the probabilistic dynamics of the system is completely defined by the equation

$$\frac{\partial S}{\partial t} = HS$$

which is the probabilistic dynamics' analogue of Schrödinger's equation. This equation has a very general solution

$$S(t) = e^{tH} S(0)$$

where the exponent of an operator is defined as

$$e^{tH} = \sum_{k=0}^{\infty} \frac{(tH)^k}{k!}$$

and eponentiation of an operator is just repeated application (taking care to respect the order of non-commutative operators).

Going back now to the agent that has propensity $r$ to annihilate itself, the equation of motion is

$$\frac{\partial S}{\partial t} = H_a S = \left( \frac{\partial}{\partial z} - z\frac{\partial}{\partial z} \right) rS$$

Solving this for $S(0) = z$ gives

$$S(t) = e^{rt((1-z)\frac{\partial}{\partial z})} z$$

$$= \sum_{k=0}^{\infty} \frac{\left( rt(1-z)\frac{\partial}{\partial z} \right)^k}{k!} z$$

$$= \sum_{k=0}^{\infty} \frac{(-rt)^k \left( (z-1)\frac{\partial}{\partial z} \right)^k}{k!} z$$

but since $\frac{\partial^n}{\partial z^n} = 0$ for $n > 1$

$$\left( (z-1)\frac{\partial}{\partial z} \right)^k = (z-1)\frac{\partial}{\partial z}$$

for $k > 0$, so

$$S(t) = 1 + \sum_{k=0}^{\infty} \frac{(-rt)^k}{k!}(z-1)$$

where the initial 1 accounts for the case when $k = 0$. We can now turn the sum back into an exponent to give

$$S(t) = (1 - e^{-rt}) + e^{-rt} z$$

So the probability that the agent exists reduces exponentially as time progresses. This is exactly what we would expect.

## 1.2 Many agents

We can now show how our operators extend very elegantly to systems with more than one agent. In this case, we represent the state that there are definitely $n$ agents in the system as $S = z^n$. A probabilistic state is therefore represented by a polynomial

$$S = \sum_{n=0}^{\infty} p_n z^n$$

where $p_n$ is the probability that there are $n$ agents, and

$$\sum_{n=0}^{\infty} p_n = 1$$

By choosing this representation, we can see immediately that our creation operator works on definite states:

$$a^\dagger z^n = z^{n+1}$$

Even better, it works on probabilistic states too. For example, suppose there's a 0.5 probability that the state is empty and a 0.5 probability that there's one agent, so $S = 0.5 + 0.5z$. In this case

$$a^\dagger(0.5 + 0.5z) = 0.5z + 0.5z^2$$

So there's now an equal probability of there being either one or two agents, i.e. the operator simultaneously adds an agent to all possible states.

If we start with an empty system (i.e. $S(0) = 1$) and add agents at a constant rate we'd expect the Hamiltonian to be given by

$$\frac{\partial S}{\partial t} = H_{a^\dagger} S = (a^\dagger - 1)rS$$

This can again be solved analytically:

$$S(t) = e^{tH} = e^{-rt(1-z)} = e^{-rt}e^{rtz}$$

$$= e^{-rt}\sum_k \frac{(rt)^k}{k!}z^k$$

so, the probability of there being $k$ agents (i.e. the coefficient of the $k^{th}$ power of $z$) is

$$c_k = \frac{(rt)^k e^{-rt}}{k!}$$

which is just the Poisson distribution, as we would expect.

Now let's try the annihilation operator on $S = 0.5z + 0.5z^2$:

$$a(0.5z + 0.5z^2) = 0.5 + z$$

We don't get $0.5 + 0.5z$ as you may have expected. Instead, for a polynomial $S = \sum_n p_n z^n$ we get

$$a\sum_n p_n z^n = \sum_n n p_n z^{n-1}$$

This behaviour can be harnessed to make a very useful operator

$$N = a^\dagger a$$

which in quantum field theory is called the number operator because it multiplies the probability of each state by the number of agents in that state

$$a^\dagger a\sum_n p_n z^n = \sum_n n p_n z^n$$

We can now understand the Hamiltonian of the annihilation operator in a new light

$$H_a = (a - a^\dagger a)r = (a^{\dagger -1} - 1)Nr$$

where $a^{\dagger -1} = z^{-1}$ is the inverse of the creation operator, which we'll call the agent-based annihilation operator. In this form, we can understand the $a^{\dagger -1}$ term as the transition operator for a single agent's propensity. The number operator $N$ has the effect of applying this operator to all the agents in the system, as if we'd executed a `foreach` loop in a computer program.

This works for any operator that represents the transition function for an agent's propensity. Because it's so useful, well define a higher order function which takes an agent action $T$ and returns the Hamiltonian that applies this propensity to all agents in a probabilistic state

$$\mathscr{H}_r(T) = (T - 1)Nr$$

we'll call this the *foreach* operator.

To illustrate this, suppose we start with $n$ agents and each agent has a propensity, $r$, to annihilate itself. In pseudocode this may look something like

```
foreach(a in S) {
  if(random() < r*dt) delete(a)
}
```

We can create the Hamiltonian for this

$$\frac{\partial}{\partial t} = \mathscr{H}(a^{\dagger -1}) = (a^{\dagger -1} - 1)Nr$$

Expanding this gives

$$H = (a^{\dagger -1} - 1)a^\dagger ar = (a - a^\dagger a)r$$

which is exactly the Hamiltonian from the single agent case, but now we understand why it has this form and why it will also work in the case of many agents.

The only difference here is the boundary condition. Instead of $S(0) = z$ as in the single agent case, we now have $S(0) = z^n$. The solution to this is

$$S(t) = \sum_m \binom{n}{m} e^{-rtm}(1 - e^{-rt})^{n-m} z^m$$

The coefficients of $z$ (i.e. the probabilities that there are $m$ agents) make up a Binomial distribution, which, again, is what we would expect.

The above examples, although very simple, demonstrate how we can define a Hamiltonian operator in terms of the behaviour of a single agent then apply it to each agent in a system and ultimately describe the aggregate dynamics of a probability distribution over the whole system.

## 1.3 Agent states

So far, the agents we've been considering haven't had any internal state. We'll start simple again by giving agents a binary state which can be either 0 or 1. We'll represent an agent in state 0 as $z_0$ and an agent in state 1 as $z_1$. A state with $n$ agents in state $z_0$ and $m$ agents in state $z_1$ is represented as state $S = z_0^n z_1^m$. We define annihilation and creation operators for each state:

$$a_0 = \frac{\partial}{\partial z_0}$$

$$a_1 = \frac{\partial}{\partial z_1}$$

$$a_0^\dagger = z_0$$

and

$$a_1^\dagger = z_1$$

Similarly, we now have two number operators $N_0 = a_0^\dagger a_0$ and $N_1 = a_1^\dagger a_1$ and two foreach operators $\mathscr{H}_{0r}(T) = (T -$

$1)N_0 r$ and $\mathscr{H}_{1r} = (T-1)N_1 r$ which loop over agents in state 0 and 1 respectively.

A change of state from $x$ to $y$ can be represented as an agent in state $x$ annihilating itself while simultaneously creating a new agent in state $y$, $a_y^\dagger a_x^{\dagger -1} = z_y z_x^{-1}$.

Let's illustrate this with a system of $n$ agents who each have a propensity $r$ to flip to the opposite state. In pseudocode we want something like

```
S0  = S
foreach(x in S0 suchthat S0.state(x)==0) {
  if(random() < r*dt) S.state(x) = 1
}
foreach(x in S0 suchthat S0.state(x)==1) {
  if(random() < r*dt) S.state(x) = 0
}
```

This can be expressed as a Hamiltonian

$$H = r\mathscr{H}_{0r}(z_1 z_0^{-1}) + r\mathscr{H}_{1r}(z_0 z_1^{-1})$$

expanding this gives

$$H = r\left((z_1 - z_0)\frac{\partial}{\partial z_0} + (z_0 - z_1)\frac{\partial}{\partial z_1}\right)$$

If we suppose that at time $t = 0$ all agents are in state $z_1$ then we can solve

$$S(t) = 2^{-n}\sum_{m=0}^{n}\binom{n}{m}(1-e^{-2rt})^{n-m}(1+e^{-2rt})^m z_0^{n-m} z_1^m$$

[TODO: add proof]

We can represent an agent with more than one variable in it's state quite naturally by allowing vectors in the subscripts of $z$. So, for example, an agent with two binary values, $a$ and $b$ could be referred to as $z_{[a,b]}$.

## 1.4 Interacting agents

Now let's get agents to interact. Suppose we have a compartmental model of infectious disease spread. Suppose there's a population of $n$ people. Each person can be one of suseptible, infected or recovered. Infected people have a propensity, $\beta$, to meet and infect a suseptible person, in which case the suseptible person becomes infected. Also, infected people have a propensity, $\gamma$ to recover.

In order to represent the interaction of infected people with suseptible, we want a Hamiltonian that would be equivalent to pseudocode something like

```
foreach(x in S0 suchthat
        S0.state(x)==Suseptible) {
  foreach(y in S0 suchthat
          S0.state(y) == Infected) {
    if(random() < beta*dt)
      S.state(x) = Infected
  }
}
```

It turns out that the number operator can be used in a quite natural way to represent this nested `foreach` operator.

If we let $z_S$, $z_I$ and $z_R$ represent suseptible, infected and recovered people respectively, then the Hamiltonian we're after is just

$$H = \mathscr{H}_{S\beta}(z_I z_S^{-1} N_I) = (z_I z_S^{-1} N_I - 1)N_S\beta$$

so, the Hamiltonian for the whole model is just

$$H = \mathscr{H}_{S\beta}(z_I z_S^{-1} N_I) + \mathscr{H}_{I\gamma}(z_R z_I^{-1})$$

and the equation of motion for the probabilistic state is, as always

$$\frac{\partial S}{\partial t} = HS$$

Although this is a differential equation, it's worth pointing out explicitly that it is representing discrete interactions between agents. So, it captures things like disease eradication when the number of infected agents is small, unlike the standard differential version of the SIR model, which is only accurate for large numbers of agents.

## 1.5 Arithmetic

If an agent has numerical variables in its state, we can represent arithmetic on those variables with a change of state operation. For example, if an agent has two integers, $a$ and $b$ in its state, we can represent the operation $a = a + b$ as

$$z_{[a+b,b]} z_{[a,b]}^{-1}$$

and so the Hamiltonian for an agent that has a unit propensity to perform this operation is

$$H = \mathscr{H}_{[a,b]}(z_{[a+b,b]} z_{[a,b]}^{-1}) = (z_{[a+b,b]} z_{[a,b]}^{-1} - 1)N_{[a,b]}$$

The same method can be used to apply any operator on any number of variables.

## 1.6 Continuous states and the uncertainty principle

So far, we've only considered agents with a finite number of discrete states. We may want to endow the agent with

a real number as part of its state, in which case the agent can be in any one of an uncountable infinity of states and we'd need a polynomial in an infinite number of variables to represent this.

The idea of our state polynomial carries over to the infinitesimal case quite naturally. The transition is made easier if we think of the coefficients as a function from monomials to coefficient values. For example, for a monovariate polynomial

$$P = \sum_n c_n z^n = \sum_n f(z^n) z^n$$

In this format we can go over to the infinitesimal limit by replacing the monomials with sets of real numbers. The coefficients can now be represented as a function from sets of real numbers to probability densities (i.e. a PDF over the set of sets of real numbers). Because we chose sets rather than bags to represent monomials, we can now no longer represent multiple agents in the same state, but because each state is infinitesimally small the probability of finding more than one agent in a state is of second order.

The construction and annihilation operators now become the set member addition and removal operators, where removal of a member from a set that does not contain that member results in the *null* set (which is distinct from the empty set) and addition to the null set results in the null set.

Consider an agent with a 1-dimensional, real valued state. We'd naturally represent a probability distribution over this value as a density function $P : \mathscr{R} \to \mathscr{R}$. For such a density function, we define the information of the function to be

$$I = \int_{-\infty}^{\infty} P(x) \log(P(x)) dx$$

We can escape the infinitude of variables by noting that in all realistic scenarios, the amount of information we have about an agent is finite. That is, whatever observations we make of a variable, there always remains some inherent uncertainty. This is analogous to Heisenberg's uncertainty principle: the more precicely we measure the location of a particle, the more momentum it must have. In our case, momentum is replaced with information and we are only interested in cases of finite information.

Given this, we can represent any probability density that has finite information as a probability mass function over a finite number of variables. For example, if the PDF has a compact support, a convenient representation would be as a Bernstein polynomial, a Deselby polynomial or the spectral modes.

[TODO: find good bases that make the mapping simple from agent to aggregate behaviour]

## 2 Making observations and doing inference

[connection to probabilistic programming]

If we have a prior $S$ and an observation operator $O$ that goes from probabilistic system states to 1D probabilistic state and we make an observation $o$ then we can define the posterior

$$P(S|o) \propto P(o|S)P(S)$$

i,e, we multiply the probability distribution over $S$ by the likelihood of the observation. We can represent likelihoods in a similar way to probabilistic states

$$S(o|S) = \sum_n P(o|S) z^n$$

However, multiplication of two probability distributions is not the same as multiplication of our state polynomials. Instead, it is multiplication of the coefficients:

$$\left(\sum_n c_n z_n\right) \times \left(\sum_m d_m z_m\right) = \sum_n c_n d_n z^n$$

Going back to the example of a single agent that has a unit propensity to annihilate itself. If we start with a prior that there's a 0.5 probability that the agent exists, then observe that it doesn't exist at time $t$, what's the posterior probability that it existed at time $t = 0$? The equation of motion is

$$\frac{\partial S}{\partial t} = HS = \left(\frac{\partial}{\partial z} - z\frac{\partial}{\partial z}\right) S$$

Integrating this for the case $S(0) = (1-p) + pz$

$$S(t) = e^{\frac{\partial}{\partial z} - z\frac{\partial}{\partial z}} S$$

$$= 1 - pe^{-rt} + e^{-rt} pz$$

So,

$$S(o|S) = 1 + (1 - e^{-rt}) z$$

So

$$S(S|o) \propto (1-p) + (1-e^{-rt}) pz$$

normalising

$$\int S(S|o) dp = (1 - 0.5e^{-rt})$$

$$S(S|o) = \frac{(1-p) + (1-e^{-rt}) pz}{(1-0.5e^{-rt})}$$

So, if the prior $S(0) = 0.5 + 0.5z$ then the posterior is given by

$$S(S|o) = \frac{0.5 + 0.5(1-e^{-rt}) z}{(1-0.5e^{-rt})}$$

# 3 Numerical methods

So far, we've considered simple systems that have analytic solutions. However, almost all of the systems of interest will not have an analytic solution. We'll now introduce some techniques to help us to numerically approximate solutions.

## 3.1 Integrating over propensities: The path integral

## 3.2 State approximation: Projecting to lower dimensional spaces

Whether or not the agents have continuous variables in their states, the number of coefficients in the probabilistic state of the system will expand exponentially as the number of variables of each agent increases. This means for many real world models we won't be able to explicitly represent the polynomial of the probabilistic state.

In order to deal with this, we will now introduce a way to project the high dimensional space of probabilistic states into lower dimensional spaces that are more practical to work with.

# 4 Data assimilation