Modelling interacting Poisson agents

Daniel Tang

May 20, 2024

Let $\mathcal{A}$ be the domain of agent states, and let an *occupation vector*, V, be an $|\mathcal{A}|$-dimensional vector of integers so that The index of each element represents a state of an agent and the integer value at that index represents the number of agents in that state.

We define a Poisson agent as an agent whose state transition probability at any time can be expressed as a set of Poisson processes with rates $\rho_1..\rho_n$ so that in infinitesimal time dt the probability of a transition due to the i^{th} process is $\rho_i dt$. Furthermore, we assume that the processes can be expressed by the following two rate functions:

1. $\rho_\psi(\Delta V)$ which is the rate at which an agent in state ψ will perform an action that results in a perturbation in occupation vector of ΔV (i.e. $V' = V + \Delta V$) and
2. $\rho_{\psi\phi}(\Delta V)$ which is the rate at which an agent in state ψ will interact with an agent in state ϕ to produce a perturbation of ΔV.

we also assume that all rates are non-negative and an agent can only be removed from the model by itself, so $\rho_\psi(\Delta V)$ can only be non-zero if all elements of $\Delta V + 1_\psi$ are non-negative, similarly $\rho_{\psi\phi}(\Delta V)$ can only be non-zero if all elements of $\Delta V + 1_\psi + 1_\phi$ are non-negative (where 1_ξ is the vector whose ξ^{th} element is one, and all other elements zero).

Given the rates of an agent's processes, we can define the Poisson rates of transition, ρ, between occupation vectors in a model containing multiple agents, so that $\rho(V'|V)dt$ is the probability that a vector V will transition to state V' in time dt. So we have a Markov model over the domain of occupation vectors whose state changes are Poisson processes whise rates are given by

$$\rho(V + \Delta V|V) = \sum_\psi \rho_\psi(\Delta V)\gamma_1(V_\psi) + \sum_{\psi,\phi} \rho_{\psi\phi}(\Delta V)\gamma_1(V_\psi)\gamma_2(V_\phi) \qquad (1)$$

where V_ψ is the value of ψ^{th} element of V, $\gamma_1(n)$ models how an agent's rate is affected by other agents in the same state, $\gamma_2(m)$ models how an agent's rate is affected when in the presence of multiple other agents in the same state. For now we consider only $\gamma_i(n) = n$ (interaction is like a chemical reaction. We'll take this to be the default behaviour if not otherwise stated), $\gamma_1(n) = (n > 0)$ (agents only care about the presence or absence of other agents, and multiplicity doesn't

affect the rate) or $\gamma_i(n) = \min(n, n_{\max})$. Note that $\gamma_i(0) = 0$ and $\gamma_i(1) = 1$ for any valid definition.

Given this, we can consider probability distributions over the domain of occupation vectors. The transition rates define a rate of change of a probability distribution,

$$\frac{d\Phi(V)}{dt} = \sum_U \rho(V|U)\Phi(U) - \sum_W \rho(W|V)\Phi(V) \tag{2}$$

In this way we can translate the agent based model into a dynamic system whose state space is the probability distributions over the domain of occupation vectors.

If we consider Φ to be a vector with an element for each occupation vector, and define ρ to be a matrix such that

$$\rho_{VU} = \begin{cases} \rho(V|U) & \text{if } U \neq V \\ \sum_W -\rho(W|V) & \text{if } U = V \end{cases}$$

then equation 2 can be expressed as the matrix equation

$$\frac{d\Phi}{dt} = \rho\Phi \tag{3}$$

so that

$$\Phi(t) = e^{\rho t}\Phi(0) = e^{(\mu - I)rt}\Phi(0) = e^{-rt}e^{\mu rt}\Phi(0)$$

where $r = \max_V -\rho_{VV}$ is a scalar and $\mu = \frac{\rho}{r} + I$ and the last step is valid since μ commutes with I.

Expanding the exponential in μ gives

$$\Phi(t) = \sum_{k=0}^{\infty} \frac{(rt)^k e^{-rt}}{k!}\mu^k\Phi(0)$$

which we recognize as a sum of powers of μ weighted with a Poisson distribution. However, since each entry of μ is non-negative and the sum of entries in each column is 1, $\mu^k\Phi(0)$ is the state of a discrete time Markov process after k steps. So, each continuous-time Poisson model ρ has an associated discrete-time Markov process $\mu = \frac{\rho}{r} + I$ such that the state of the Poisson model at time t is the weighted sum of states in a trajectory of the Markov process, where the weights are given by a Poisson distribution with rate rt.

This means, among other things, that if the Markov model has an attractor then as $t \to \infty$ the state distribution of the corresponding Poisson model tends to an equilibrium distribution, and that distribution is the uniform probability over the attractor of its Markov model. So every Poisson model tends to a steady state distribution (i.e. a single point in distribution space).

The occupation vectors can be thought of as points in an $|\mathcal{A}|$-dimensional space, and Φ can be thought of as a probability field over an A-dimensional grid on the non-negtive integer coordinate points in this space. With this in mind, we can re-express the multiplication by μ in terms of a convolution.

$$\mu\Phi = \left(\frac{\rho}{r} + I\right)\Phi = \frac{1}{r}\rho\Phi + \Phi$$

but

$$\rho\Phi = \sum_{V'} \left(\sum_{\psi} \rho_{\psi}(V - V')\gamma_1(V'_{\psi}) + \sum_{\psi,\phi} \rho_{\psi\phi}(V' - V)\gamma_1(V'_{\psi})\gamma_2(V'_{\phi}) \right) \Phi(V') \\ - \left(\sum_{\Delta V,\psi} \rho_{\psi}(\Delta V)\gamma_1(V_{\psi}) + \sum_{\Delta V,\psi,\phi} \rho_{\psi\phi}(\Delta V)\gamma_1(V_{\psi})\gamma_2(V_{\phi}) \right) \Phi(V) \tag{4}$$

Substituting 1 into 2 gives

$$\frac{d\Phi(V)}{dt} = \sum_{V'} \left(\sum_{\psi} \rho_{\psi}(V - V')\gamma_1(V'_{\psi}) + \sum_{\psi,\phi} \rho_{\psi\phi}(V' - V)\gamma_1(V'_{\psi})\gamma_2(V'_{\phi}) \right) \Phi(V') \\ - \left(\sum_{\Delta V,\psi} \rho_{\psi}(\Delta V)\gamma_1(V_{\psi}) + \sum_{\Delta V,\psi,\phi} \rho_{\psi\phi}(\Delta V)\gamma_1(V_{\psi})\gamma_2(V_{\phi}) \right) \Phi(V) \tag{5}$$

---

simplify equation 5 by expressing the terms as convolutions. Let the "transition kernels" be

$$\tau_{\psi}(\Delta V) = \begin{cases} \rho_{\psi}(\Delta V) & \text{if } \Delta V \neq 0 \\ \sum_{\Delta V'} -\rho_{\psi}(\Delta V') & \text{if } \Delta V = 0 \end{cases}$$

and

$$\tau_{\psi\phi}(\Delta V) = \begin{cases} \rho_{\psi\phi}(\Delta V) & \text{if } \Delta V \neq 0 \\ \sum_{\Delta V'} -\rho_{\psi\phi}(\Delta V') & \text{if } \Delta V = 0 \end{cases}$$

and let $n_{i\psi}$ denote an A-dimensional field such that $n_{i\psi}(V) = \gamma_i(V_{\psi})$ so that the sums over V' and ΔV become convolutions

$$\frac{d\Phi}{dt} = \sum_{\psi} \tau_{\psi} * n_{1\psi}\Phi + \sum_{\psi,\phi} \tau_{\psi\phi} * n_{1\psi}n_{2\phi}\Phi \tag{6}$$

where multiplication of fields is to be interpreted as being pointwise.

We can calculate higher order rates of change, since Φ is the only term with time dependency we have

$$\frac{d^{n+1}\Phi}{dt^{n+1}} = \sum_{\psi} \tau_{\psi} * n_{1\psi} \frac{d^n\Phi}{dt^n} + \sum_{\psi,\phi} \tau_{\psi\phi} * n_{1\psi}n_{2\phi} \frac{d^n\Phi}{dt^n} \tag{7}$$

---

1 Public states and Interaction states

We can distinguish between different classes of interaction rate function $\rho_{\psi\phi}$ in order to understand how the properties of this function affect the ABMs dynamics, and also to make computationally efficient algorithms.

It may be possible to split the state of an agent into public and private states, $\psi = \langle \psi_u, \psi_l \rangle$ so that the interaction rate depends only on the public state of the other agent $\rho_{\psi\langle\phi_u,\phi_l\rangle} = \rho'_{\psi\phi_u}$, thereby reducing the dimensionality. Occupation vectors can then be represented as matrices where the row gives the public state and the column gives the private state.

A speical case of this is when only agents with the same public state can interact. In this case, we call the public state an "interaction state".

2 Simulating

Given an occupation vector and a set of transitions with associated rates, the probability of a transition along edge i in time dt is $\rho_i dt$ and the probability that there is no transition in dt is $1 - \sum_k \rho_k dt$, so the probability that i is the next transition is

$$P_n(i) = \sum_{t=0}^{\infty} (1 - \sum_k \rho_k dt)^t \rho_i dt = \frac{\rho_i}{\sum_k \rho_k}$$

so we just need to choose a transition with probability proportional to the rate.

If we express transitions as perturbations to the occupation vector, then on transition most transitions remain unchanged, so we can implement this as a distribution of transitions from which we choose a transition and use the chosen transition to update the dsitribution.

3 Spatial agents

Suppose agents are in a spatial environment (e.g. a 2D grid) and that only agents on the same grid-square can interact. Within a gridsquare, all pairwise interactions are possible, but the rates are dependent on the non-spatial internal states of the agents. An event can consist of a single method call, or a whole turns-based encounter.

If an event is just a method call, then even binary events change only the state of a single agent. which may simplify parallelisation. In this case every agent has a local time set to the time of its last state-change. The time of its next state change is dependent only on the rates of its own state-changing processes which are defined by the states of the other agents in the same grid-square at the agent's current local time. So, a grid-square needs to "remember" the states of all agents at each of the local times of the contained agents. An agent cannot step forward until the rates of all its interactions are known.

If an event is a whole binary encounter, then we have improved computational efficiency (more useful computation per event, no need to remember who we're talking with inbetween events, no need for high rate responses). However, if encounters are non-deterministic, given the full state of both agents, then we need to encode probabilistic consequences of an event (this could be encoded by simulation of the encounter). If agents are identifiable, an agent can have high rates of interaction with known others.

In either case, a binary interaction has a primary and a secondary agent (i.e. an interaction is not symmetrical, one of the agents initiates the interaction, so an interaction of A and B is not necessarily the same as an interaction of B and A). It is also up to the initiator to set the rate of the interaction i.e. the rate of a binary interaction of A and B can depend on the complete state of A but only the public state of B.

3.1 Parallelism in spatial ABMs

Suppose we assign local times to grid-squares so that all agents within a grid-squre are at the same gridsquare local time. Each gridsquare is computed with at most one thread. We specify that it takes time Δt for an agent to move from one gridsquare to another, during which time no other events may occur to the agent. So a gridsquare can calculate up to its "frontier" which is Δt ahead of the earliest local time of its neighbouring squares. In this way we get a "light cone" constraint.

Each gridsquare has a local time, t, a frontier time, t_f, a set of neighbours that this square is currently blocking and an "incoming agent" list. If $t = t_f$ we say the square is blocked. After an event is processed for a gridsquare, a "next event" is drawn for the interval $[t, t_f]$. If the set of Poisson processes contained in the gridsquare have rates $\rho_1 ... \rho_N$ and $R = \sum_i \rho_i$ then the probability that the next event occurs in the interval $[\tau, \tau + d\tau]$ from the current time is given by

$$P_i(\tau)d\tau = Re^{-R\tau}d\tau$$

So, draw a time offset τ from this exponential, if $t + \tau > t_f$ then the gridsquare becomes blocked on the earliest neighbour (randomly chosen if there are multiple earliest neighbours). Otherwise, $t \to t + \tau$, a process is chosen with probability $\frac{\rho_i}{R}$ and the event is added to a set of events to be processed.

If the gridsquare's incoming-agent list is not empty and the next incoming agent is at time $t < t' < t_f$ then we treat the interval $[t, t']$, add the agent if no intervening event occurs, otherwise process the event, and repeat.

When a gridsquare's local-time is updated, any neighbours that are blocked on this one are notified of the new local time. When a blocked gridsquare receives an unblocking notification, it updates its frontier and draws an event as above, possibly blocking again and possibly sending further unblocking notifications.

A simulation begins with all gridsquares at $t = 0$ and all frontiers at $t_f = \Delta t$. Each gridsquare is drawn from to create the initial set of events to be processed, so that all squares are either blocked or have an event to be processed.

If an agent moves from gridsquare i to gridsquare j, then the agent is removed

from i and added to the destination gridsquare's incoming-agent list with timestamp $t_i + \Delta t$.

4 Computing over distributions

4.1 Computing with non-local basis

5 Computation for studying social norms

[What do we want to calculate? How should we use computation to do sociology?

Existance proof that an observed social phenomenon does not require a given individual property. Proved by simulation of individuals that lack the given property yet robustly (i.e. not by luck/chance) demonstrate the given observed social behaviour.

Calculation of posterior over individual behaviour given a prior and a set of social observations, in the form of timeseries or of properties of the attractor. Can be used to imply or engineer individual behaviour.

Data assimilation, given individual behaviour and timeseries observation (or perhaps only prior over behaviour).

Social engineering over social norms: Given a space of social norms, an individual behaviour (prior?) and an objective function, find a set of social norms that maximises the expectation of the objective function on the attractor [do societies ever reach the attractor? Can we expect all point on an attractor to have the same social norms (i.e. is it a point in social norm space)?].

Stability analysis: Given a set of social norms, how stable is this set to deviant behaviour, subcultures, endogenous evolution, exogenous forcing?

Engineering Revolution: Given a set of social norms, find the smallest (a small) perturbation, or set of small perturbations that leads to a new set of norms.

]