The compulsive gambler process

David Aldous

March 17, 2015

Liggett's 1984 monograph *Interacting Particle Systems* was very influential in introducing to the broad mathematical probability community a class of models, originating in statistical physics, exemplified by the Ising model on $\mathbb{Z}^3$. Such models are now used, and often reinvented, in many disciplines outside physics, in particular in the context of social networks. The physics setting suggests

- infinite, highly structured network of sites ($\mathbb{Z}^3$)
- categorical states (± 1)
- study phase transitions for the equilibrium distribution

whereas the social networks setting suggests

- finite, unstructured network of agents
- allow numerical states ($\mathbb{R}$)
- study finite-time behavior (cf. "Markov Chains and Mixing Times").

In my 2013 survey *Interacting Particle Systems as Stochastic Social Dynamics* I formalize this viewpoint as a **FMIE process;**

- n agents; each in some state $\in S$
- each pair of agents (i, j) meet at times of a Poisson process of given rate ν_{ij}
- agents enter such a meeting in states $(X_i(t-), X_j(t-))$, leave in states $(X_i(t+), X_j(t+))$ given by some deterministic or random update rule $F : S \times S \to S \times S$.

The update rule specifies the process, e.g. the familiar **voter model**.

The process makes sense for any "meeting model" specified by the (ν_{ij}), best viewed as an arbitrary connected finite edge-weighted graph.

In this talk I briefly discuss a simple example, the **averaging process**, before getting to the main topic. In both models, "state" is most naturally interpreted as money. What follows is work with (or by) graduate student Dan Lanoue, plus contributions from Justin Salez.

Model: Averaging Process

When agents i and j meet, they split their combined money equally, so the values ($X_i(t)$ and $X_j(t)$) are replaced by the average $(X_i(t) + X_j(t))/2$.

The overall average is conserved, and obviously each agent's fortune $X_i(t)$ will converge to the overall average. Note a simple relation with the associated continuous-time Markov chain. Write 1_i for the initial configuration $X_j(0) = 1_{(i=j)}$ and $p_{ij}(t)$ for transition probabilities for the Markov chain.

Lemma

In the averaging process started from 1_i we have $\mathbb{E}X_j(t) = p_{ij}(t/2)$. More generally, from any initial configuration $\mathbf{x}(0)$, the vector of expectations $\mathbf{x}(t) := \mathbb{E}\mathbf{X}(t)$ evolves exactly as the dynamical system

$$\frac{d}{dt}\mathbf{x}(t) = \frac{1}{2}\mathbf{x}(t)\mathbb{N}$$

where $\mathbb{N}$ is the generator of the associated MC.

So if $\mathbf{x}(0)$ is a probability distribution, then the means evolve as the distribution of the MC started with $\mathbf{x}(0)$ and slowed down by factor $1/2$.

It turns out to be easy to quantify global convergence to the average.

Proposition (Global convergence in the Averaging Process)

From an initial configuration $\mathbf{x} = (x_i)$ *with average zero and* L^2 *size* $||\mathbf{x}||_2 := \sqrt{n^{-1} \sum_i x_i^2}$, *the time-t configuration* $\mathbf{X}(t)$ *satisfies*

$$\mathbb{E}||\mathbf{X}(t)||_2 \leq ||\mathbf{x}||_2 \exp(-\lambda t/4), \quad 0 \leq t < \infty \tag{1}$$

where λ *is the spectral gap of the associated MC.*

Results like this have appeared in several contexts, e.g. gossip algorithms.

Here is a more subtle result. Suppose normalized meeting rates:

$$\sum_{j \neq i} \nu_{ij} = 1 \ \forall i.$$

Because an agent interacts with nearby agents, guess that some sort of "local averaging" occurs independent of the geometry.

For a "test function" g : **Agents** $\to \mathbb{R}$ write

$$\begin{aligned} \bar{g} &= n^{-1} \sum_i g_i \\ ||g||_2^2 &= n^{-1} \sum_i g_i^2 \\ \mathcal{E}(g,g) &= n^{-1} \tfrac{1}{2} \sum_i \sum_{j \neq i} \nu_{ij} (g_j - g_i)^2 \quad \text{(the Dirichlet form).} \end{aligned}$$

When $\bar{g} = 0$ then $||g||_2$ measures "global" variability of g whereas $\mathcal{E}(g,g)$ measures "local" variability relative to the underlying geometry.

Proposition (Local smoothness in the Averaging Process)

For normalized meeting rates associated with a r-regular graph; and initial $\bar{\mathbf{x}} = 0$,

$$\mathbb{E} \int_0^\infty \mathcal{E}(\mathbf{X}(t), \mathbf{X}(t)) \, dt = 2||\mathbf{x}||_2^2. \quad (2)$$

The compulsive gambler process

In the **Compulsive Gambler** process, agents initially have 1 unit money (visualize as a 5-pound note) each. When two agents with non-zero money meet, they instantly play a fair game in which one wins the other's money.

This is a "made up" model – invented in preparing lectures for 2012 Cornell summer school. But interesting as pedagogy

- 4 different techniques are useful for studying this process.
- Intriguing analogies with other probability topics.

This loosely fits the theme of the workshop – obvious tree structure of solvent (non-zero money) agents as time increases.

Recall the process is parametrized by the meeting rates (ν_{ij}) – visualize as finite edge-weighted graph.

First observation: on a complete graph, that is if

$$\nu_* := \min_{j \neq i} \nu_{ij} > 0 \quad (*)$$

then at some random time $T < \infty$ one agent has all the money.

In general (if (*) fails) then process gets absorbed in some random configuration where the set of solvent agents is an anti-clique (independent set).

Methodology 1: Comparison with the Kingman Coalescent chain (which is the mean-field model $\nu_{ij} \equiv 1$). By considering number of agents with non-zero money

$$2(1 - \tfrac{1}{n})/\nu^* \leq \mathbb{E}T \leq 2(1 - \tfrac{1}{n})/\nu_*$$

where $\nu^* = \max_{j \neq i} \nu_{ij}$.

Methodology 2. Martingale properties

Write $X_i(t)$ = amount of money for agent i at time t.

Lemma

For any meeting process,
(a) $(X_i(t), 0 \leq t < \infty)$ *is a martingale.*
(b) For $j \neq i$, $(X_i(t)X_j(t), 0 \leq t < \infty)$ *is a supermartingale.*
(c) For f : **Agents** $\to \mathbb{R}$ *write* $S_f(t) = n^{-1} \sum_i f(i) X_i(t)$.
Then $(S_f(t), 0 \leq t < \infty)$ *is a martingale and*

$$\mathbb{E} S_f^2(t) - S_f^2(0) \leq \nu^* t \ \mathrm{var}[f(\xi)]$$

where ξ *is uniform random on* **Agents**.

A slightly more subtle observation is

Lemma

Let $\psi : \textbf{Agents} \times \textbf{Agents} \to [0, \infty)$ *be such that* $\psi(i,j) \equiv \psi(j,i)$ *and* $\psi(i,i) \equiv 0$. *Define*

$$\Psi(\mathbf{x}) = \sum_{\{i,j\}} x_i x_j \psi(i,j), \quad \Psi_\nu(\mathbf{x}) = \sum_{\{i,j\}} \nu_{ij} x_i x_j \psi(i,j).$$

Then the process

$$\Psi(\mathbf{X}(t)) + \int_0^t \Psi_\nu(\mathbf{X}(s)) \, ds$$

is a martingale.

Setting $\psi(i,j) = 1/\nu_{ij}$ and applying OST at the coalescence time T leads to a modest improvement on the "obvious" lower bound from the Kingman coalescent, as follows

$$\frac{1}{\nu^*} \le \frac{1}{\binom{n}{2}} \sum_{\{i,j\}} \frac{1}{\nu_{ij}} \le \mathbb{E} T. \quad (3)$$

Methodology 3. Imagine the initial currency notes have IID random serial numbers. The Compulsive Gambler process has the same distribution (unconditionally, if you don't see the serial numbers) as the process in which the winner of each bet is determined **deterministically** as the possessor of the note with the lowest serial number.

(Proof by careful induction).

(Live demo).

This means we can construct the CG process via the following **token process** representation.

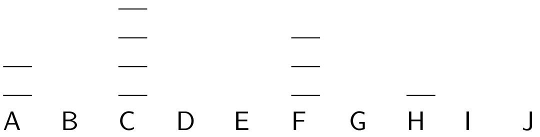
Take the array (T_{ij}) of the Exponential(rate ν_{ij}) first meeting times. Take a uniform random ordering $i_1, \ldots, i_n$ of the n agents, representing the original possessors of the notes (1), (2), . . . , (n) in increasing order of serial number.

- note (1) stays with agent i_1 for ever
- note (2) stays with agent i_2 until time $T_{i_2 i_1}$, then passes to agent i_1
- note (3) starts at agent i_3; will pass to agent i_2 at time $T_{i_3 i_2}$ if this time is smaller than $\min(T_{i_3 i_1}, T_{i_2 i_1})$, otherwise pass to agent i_1 at time $T_{i_3 i_1}$
- . . . . . .
- . . . . . .

and finally $X_i(t) =$ number of notes agent i has at time t.

Reminiscent of other constructions, e.g. Donnelly-Kurtz "look down" in math population genetics.

Methodology 4. Using the token process representation above, given $\mathbf{X}(t)$ – the amount of money each agent has at time t



then the allocation of the ordered serial numbers $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10 = n)$ is uniform random on the set of compatible partitions.

This is an exchangeability property, reminiscent of the theory of exchangeable coalescents (Bertoin et al). But it's exchangeability of currency notes, not of Agents.

What can we learn using these tools? Using either the martingale property or the token process representation

> *(if $\nu_* > 0$) the agent who ultimately acquires all the money is uniform random over all n agents.*

A more subtle fact is that, for arbitrary f : **Agents** $\to \mathbb{R}$, there is a formula for the k'th moment

$$\mathbb{E}\left(\sum_i f(i) X_i(t)\right)^k .$$

The formula consists of summing, over k-tuples $(v_1, \ldots, v_k)$ of agents, a function $p(v_1, \ldots, v_k; t)$ which consists of a sum of M_k terms, each term being of the form

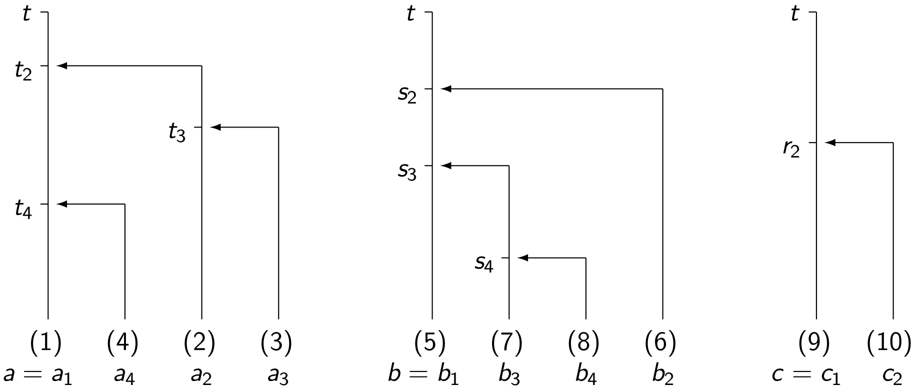
$$\prod_{i=2}^{k} \nu(v_{\alpha_i} v_i) \int_D \exp\left(-\sum_{i=2}^{k}\sum_{j=1}^{k-1} \nu(v_{\beta_{ij}} v_i) t_i\right) dt_2 dt_3 \ldots dt_k \qquad (4)$$

where $\alpha_i < i$ and $\beta_{ij} < i$ are indices and D is a subset of $[0, t]^{k-1}$, specified by the particular term.

To outline the argument, consider $k = 10$'th moment for large n. We need to understand where k-tuples of notes are; then by the exchangeability property we can reduce to a calculation with the k lowest-ranked notes (1),. . . , (k). Consider a typical possibility

agent a has $\{1, 2, 3, 4\}$, agent b has $\{5, 6, 7, 8\}$, agent c has $\{9, 10\}$.

We can find this probability from the structure of the "serial number" process.


t
t_2
t_3
t_4
(1) (4) (2) (3)
$a = a_1$ a_4 a_2 a_3
t
s_2
s_3
s_4
(5) (7) (8) (6)
$b = b_1$ b_3 b_4 b_2
t
r_2
(9) (10)
$c = c_1$ c_2

These are **tools** – where should we look to find some more substantive **Theorems**?

One direction is to consider the "sparse graph limits" setting. Take a r-regular n-vertex connected graph G; set meeting rates $\nu_e = 1/r$ for each edge e. There is some limit density of solvent agents

$$a(G) := n^{-1} \sum_i \mathbb{P}(X_i(\infty) \neq 0).$$

Consider limits of graphs as $n \to \infty$ with r fixed:

Question; what is the range $[a_*(r), a^*(r)]$ of possible limits of $a(G)$?

Partial answer; as $r \to \infty$,

$$a_*(r) \sim 1/r; \quad a^*(r) \geq 2 - o(1))/r$$

and we conjecture $a^*(r) \sim 2/r$.
The fact $a_*(r) \sim 1/r$ is easy; the inequality for $a^*(r)$ comes from considering the r-regular infinite tree $\mathbb{T}_r$ as the local limit of suitable r-regular graphs.

A conceptually interesting feature of the CG process is that one can consider both types of $n \to \infty$ process; infinite discrete space or a scaling limit on a continuous space. Regarding infinite discrete space we can consider:

the r-regular infinite tree $\mathbb{T}_r$ with rate-1 meetings across edges.

This can be analyzed by setting up a recursion to obtain an equation for the generating function $\phi(z,t)$ of $X(t)$, the fortune of a typical agent. We can estimate $\mathbb{P}(X(\infty) > 0)$ as a function of r well enough to show it is $\sim 2/r$ as $r \to \infty$.

Cute aside: on the Poisson(c)-Galton-Watson tree $\mathbb{T}_c$ one can solve the recursion to show that

$$\mathbb{P}(X(t) > 0) = \tfrac{2}{2+cs(t)}; \quad s(t) = 1 - e^{-t}$$

the conditional dist. of $X(t)$ given $X(t) > 0$ is Geometric($\tfrac{2}{2+cs(t)}$).

This process is in fact the short-time limit of the Kingman coalescent – regarding the latter as coalescence of n atoms, then the time-τ/n limit distribution of typical cluster size is Geometric($\tfrac{2}{2+\tau}$).

Another case is $\mathbb{Z}^d$ with meeting rates are

$$\nu_{ij} = ||j - i||^{-\alpha}$$

for some $\alpha > d$. Consider the mean density of solvent agents at time t

$$\rho(t) := \mathbb{P}(X_i(t) \neq 0)$$

and the conditional distribution $X^*(t)$ of $X_i(t)$ given $X_i(t) \neq 0$, for which $\mathbb{E}X^*(t) = 1/\rho(t)$ because $\mathbb{E}X_i(t) \equiv 1$. Heuristic arguments, based on supposing the positions of solvent agents do not become "clustered", suggest that

$$\rho(t) \asymp t^{-\beta} \text{ for } \beta = \tfrac{d}{\alpha}.$$

We conjecture that

$$\rho(t)X^*(t) \to_d X^*, \quad \text{for some } X^* \text{ such that } \mathbb{E}X^* = 1$$

and then that the process has a scaling limit, the limit being a process whose states are (locally finite support) measures on $\mathbb{R}^d$.

Consider the same meeting rates

$$\nu_{ij} = ||j - i||^{-\alpha} \tag{5}$$

but now on the finite discrete torus $\mathbb{Z}_m^d$ and take meeting rates as at (5). This process now has qualitatively different behavior in the case $\alpha > d$ (agents mostly meet nearby agents) and in the case $\alpha < d$ (agents mostly meet distant agents). Somewhat surprisingly we can establish the order of magnitude of T in both cases without any detailed analysis: the "harmonic mean" inequality (3) easily implies that there exist constants $c_{d,\alpha}$ and $C_{d,\alpha}$ such that

$$c_{d,\alpha} m^{\alpha} \leq \mathbb{E}T \leq C_{d,\alpha} m^{\alpha}.$$

Aside. For many years I have suggested studying IPS models with rates (5); now a detailed study of phase transitions in the first-passage percolation model with those rates has been given in Chatterjee - Dey *Multiple phase transitions in long-range first-passage percolation on square lattices*.

The Metric Coalescent

Work by Dan Lanoue – major part of his Ph.D. thesis.

First do a trivial reformulation of the CG process. Rescale so there is unit money in total; initially distributed arbitrarily amongst the n agents.

Second: instead of visualizing the n agents as vertices of a graph, we place them at positions $s_1, \ldots, s_n$ in a metric space (S, d). Define meeting rates as a function of distance

$$\nu_{ij} = \phi(d(s_i, s_j)) > 0$$

$$d \to \phi(d) \text{ continuous decreasing; } \lim_{d\to\infty} \phi(d) = 0, \lim_{d\to 0} \phi(d) = \infty.$$

Can now regard the state space of this CG process as the space $\mathcal{P}_{\mathrm{fs}}(S)$ of probability measures on S with finite support.

Write $\mathcal{P}(S)$ = space of all probability measures on S, with topology of weak convergence.

Given a general (in particular, a non-atomic) distribution $\mu \in \mathcal{P}(S)$, choose $(s_i, 1 \leq i < \infty)$ such that

$$\mu^{(n)}(0) := \text{ uniform dist. on } (s_1, \ldots, s_n) \quad \to \mu.$$

Natural to guess that the CG processes $(\mu^{(n)}(t), 0 \leq t < \infty)$ started from $\mu^{(n)}(0)$ converge (as $n \to \infty$) to some limit process, which at times $t > 0$ has locally finite support but which converges to μ as $t \downarrow 0$.

Theorem

Under suitable regularity conditions on (S, d) *and* ϕ*:*
For each $\mu(0) \in \mathcal{P}(S)$ *there exists a unique (in distribution)* $\mathcal{P}(S)$*-valued process* $(\mu(t), 0 \le t < \infty)$ *specified by*

- $(\mu(t), t_0 \le t < \infty)$ *evolves as the CG process (each* $t_0 > 0$*)*
- $\mu(t) \to \mu(0)$ *a.s. as* $t \downarrow 0$.

And the Feller property holds: for fixed $t > 0$ *the map* $\mu(0) \to \mathrm{dist}(\mu(t))$ *is continuous.*

Formally, this is describing the entrance boundary of the metric-space-embedded CG. We call the process the **metric coalescent**. The current paper proves this under the condition

(S, d) *is separable and locally compact.*

Proof is technical in detail, but uses in part the methodologies described earlier.

'[repeat 2 previous slides]

Methodology 3. Imagine the initial currency notes have IID random serial numbers. The Compulsive Gambler process has the same distribution (unconditionally, if you don't see the serial numbers) as the process in which the winner of each bet is determined as the possessor of the note with the lowest serial number.

(Proof by careful induction).

This means we can construct the CG process as follows.

Take the array (T_{ij}) of the Exponential(rate ν_{ij}) first meeting times. Take a uniform random ordering $i_1, \ldots, i_n$ of the n agents, representing the original possessors of the notes (1), (2), ..., (n) in increasing order of serial number.

- note (1) stays with agent i_1 for ever
- note (2) stays with agent i_2 until time $T_{i_2 i_1}$, then passes to agent i_1
- note (3) starts at agent i_3; will pass to agent i_2 at time $T_{i_3 i_2}$ if this time is smaller than $\min(T_{i_3 i_1}, T_{i_2 i_1})$, otherwise pass to agent i_1 at time $T_{i_3 i_1}$
- ......

Now in our metric space setting, do the same construction (with 1 unit notes) using IID samples $s_1, s_2, \ldots$ from a given $\mu(0)$. Then for each n set $\bar{\mu}^{(n)}(t) =$ the configuration of money at t, rescaled so total money $= 1$. This provides an "a.s." construction, in the sense that for fixed $t > 0$

$$\bar{\mu}^{(n)}(t) \to \text{ some } \bar{\mu}(t) \text{ a.s. as } n \to \infty.$$

Proved by a (slightly subtle) exchangeability argument (then de Finetti).

[repeat another earlier slide]

A more subtle fact is that, for arbitrary f : **Agents** $\to \mathbb{R}$, there is a formula for the k'th moment

$$\mathbb{E}\left(\sum_i f(i) X_i(t)\right)^k$$

which involves summing, over k-tuples $(v_1, \ldots, v_k)$ of agents, a function $p(v_1, \ldots, v_k; t)$ which consists of a sum of M_k terms, each term being of the form

$$\prod_{i=2}^{k} \nu(v_{\alpha_i} v_i) \int_D \exp\left(-\sum_{i=2}^{k}\sum_{j=1}^{k-1} \nu(v_{\beta_{ij}} v_i) t_i\right) dt_2 dt_3 \ldots dt_k \qquad (6)$$

where $\alpha_i < i$ and $\beta_{ij} < i$ are indices and D is a subset of $[0, t]^{k-1}$, specified by the particular term.

Key point: the number of terms in (6) depends on k, not on n

Take bounded continuous $f : S \to \mathbb{R}$; we can calculate the moments

$$\mathbb{E}\left(\int f(s)\ \mu(t, ds)\right)^k$$

in the metric space embedding as $n \to \infty$ limits of the finite-n formula; because the sum over distinct ordered k-tuples from $(s_1, \ldots, s_n)$

$$n^{-k} \sum_{v_1} \ldots \sum_{v_k} p(v_1, \ldots, v_k; t)$$

converges to the limit

$$\int \ldots \int p(s_1, \ldots, s_k; t)\ \mu(ds_1) \ldots \mu(ds_k)$$

for $\mu = \mu(0)$.

Setting aside details, we have two separate arguments for existence of a process $(\mu(t), 0 < t < \infty)$ associated with given $\mu(0) \in \mathcal{P}(S)$. The property

- $\mu(t) \to \mu(0)$ a.s. as $t \downarrow 0$.

follows from a martingale argument.

Relevant papers on arXiv

- Interacting particle systems as stochastic social dynamics (Aldous)
- A lecture on the averaging process (Aldous - Lanoue)
- The compulsive gambler process (Aldous - Lanoue - Salez)
- The Metric Coalescent (Lanoue)
- The iPod model (Lanoue)