Voter models and variants

Ayalvadi Ganesh

University of Bristol

Models

Classical voter model

- Population of n voters
- Each voter has opinion/ preference in $\{0,1\}$
- Voters interact and update their opinions …
- … in either discrete or continuous time

Classical voter model: discrete time

- Population of n voters
- Each voter has opinion/ preference in $\{0,1\}$
- Voters update their opinions synchronously
- In each time step, each voter samples a voter uniformly at random (including itself) …
- … and copies the opinion of the chosen voter

Classical voter model: continuous time

- Population of n voters
- Each voter has opinion/ preference in $\{0,1\}$
- Voters update their opinions asynchronously
- Each voter has an alarm clock that rings after independent $Exp(1)$ times, independent of the clocks at other voters
- When a voter’s clock rings, she contacts a voter chosen uniformly at random …
- … and copies the opinion of that voter

Consensus

- Voter model is a Markov process
- All-0 and all-1 states are absorbing, reachable from all other states.
- Hence, voters reach consensus eventually.

- Q: How long does it take to reach consensus?
- Q: What is the probability of reaching consensus on 1, for a given initial state?

Application: Population Genetics

- Population of n alleles
- Each allele is of one of two types, a or A
- Population size stays constant over time
- Composition evolves over time exactly as in voter model:
  - Discrete-time version is called Wright-Fisher model
  - Continuous-time version is called Moran model
- "Neutral" models – neither allele has a selective advantage over the other

Moran model in population genetics

- Population of n alleles
- Each allele is of one of two types, a or A
- Each allele dies after an $Exp(1)$ lifetime, and is replaced by a copy of another allele, chosen uniformly at random from the population

- *Q: Starting with $n-1$ "wild-type" a alleles, and a single "mutant" A allele, what is the probability that the mutant reaches fixation?*
- *Q: How long does it take until the mutant reaches either fixation or extinction?*

Moran model with selection

- Population of n alleles, of two types, a or A
- Mutant alleles A have fitness $1 + s$, wild-type a have fitness 1
- Each allele dies after an $Exp(1)$ lifetime, and is replaced by a copy of another allele, chosen at random weighted by fitness

- *Q: Starting with $n - 1$ wild-type a alleles, and a single mutant A allele, what is the probability that the mutant reaches fixation?*
- *Q: How long does it take until the mutant reaches either fixation or extinction?*

Infinite alleles Moran model

- Population of n alleles
- Initially, all alleles are of type 0
- Each allele dies after an $Exp(1)$ lifetime, and:
  - with probability $1 - \varepsilon$, it is replaced by an allele chosen uniformly at random from the population
  - with probability ε, it is replaced by a mutant of a completely new type, i.e., no two mutations give rise to the same allele

- *Q: What is the population structure in equilibrium?*

Voter model on a graph

- $G = (V, E)$: connected, undirected graph
- Nodes correspond to voters
- Voters update their opinions in discrete or continuous time, sampling uniformly from among their neighbours in G …
- … or non-uniformly according to a given stochastic matrix P (discrete time) or rate matrix Q (continuous time)
- Remark: If $G = K_n$, we recover the classical model.

- *Q: What can we say about consensus times and probabilities?*

Example: Voter model on lattices

- Infinite lattice in d dimensions
- One of two initial preferences, 0 or 1, at each vertex

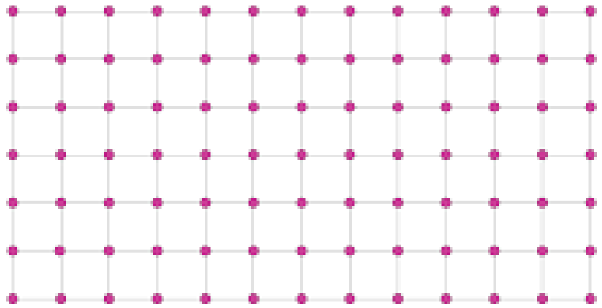


Image from Wikimedia Commons

- *Q: What happens in the long run?*

Example: Voter model on lattices

- Infinite lattice in d dimensions
- One of two initial preferences, 0 or 1, at each vertex

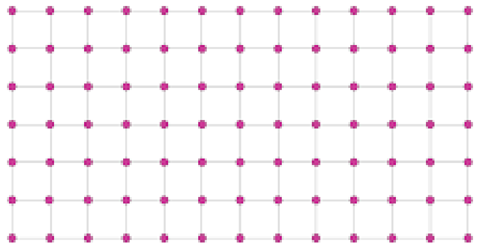


Image from Wikimedia Commons

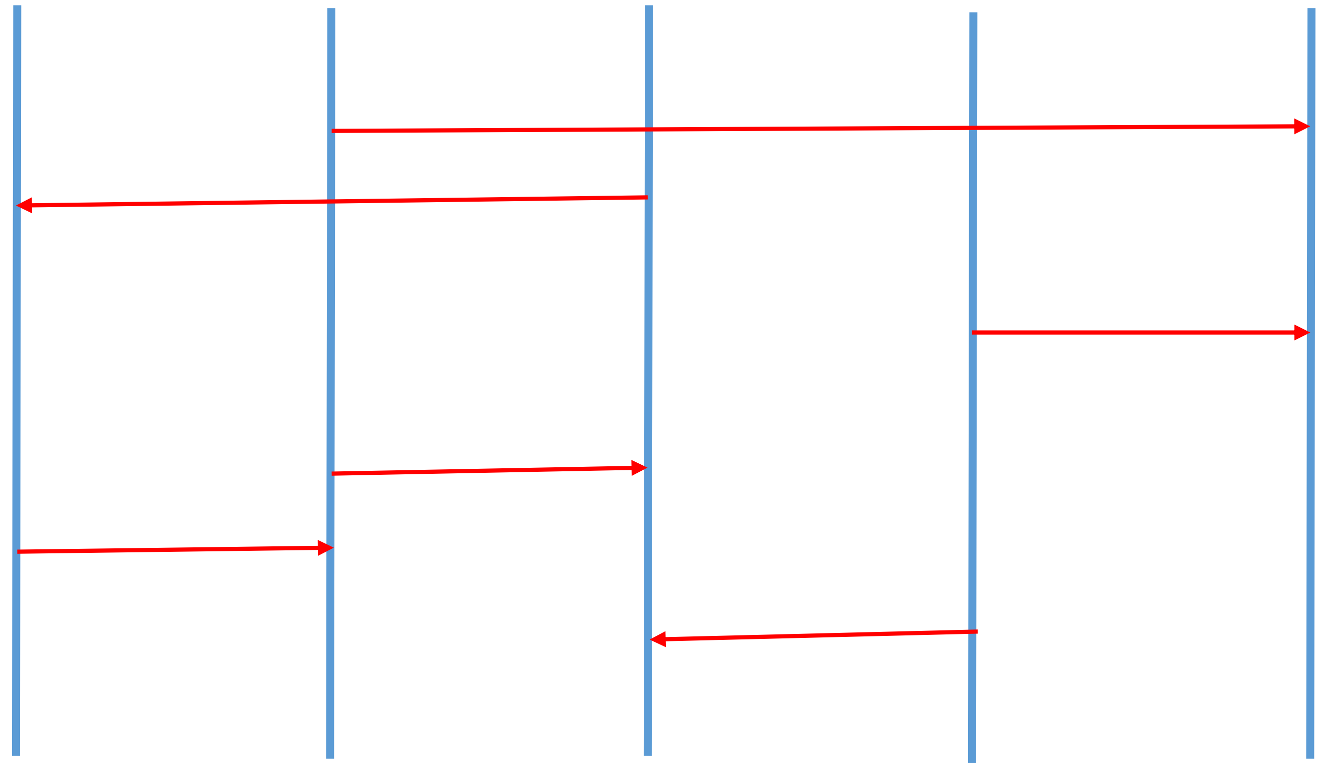
- Answer: In the long run, there is
  - consensus in dimension 1 or 2
  - co-existence in 3 or more dimensions

Applications

- Spin glasses
  - voter preference represents spin
- Evolutionary biology / population genetics
- Ecology
  - behaviour of social insects, bioluminescent bacteria (quorum sensing), flocking of fish and birds
- Economics and social science
  - competition between products and technologies, adoption of customs and behaviours, fads and fashions, movement of crowds
- Computer science
  - decentralised algorithms for averaging, consensus, information fusion

Analysis

Graphical representation and duality



Coalescing random walks

- Start with a single particle at each vertex $v \in V$
- Particles perform independent random walks with rates $q(v, w)$
- If a particle moves to an occupied site, it immediately merges with the particle there – the new particle continues to perform a random walk with the same rate

- Voter model in reversed time corresponds to a coalescing random walk

Implications

- Time to consensus in voter model is bounded by the time to have a single particle in the coalescing random walk
  - in particular, for $G = K_n$, we have $E[T] \leq n$
- Voter model in 1 and 2 dimensions achieves consensus, but in 3 or more dimensions, there is co-existence

Consensus value

- Assumption: the matrix Q or P governing the random walk admits a unique invariant distribution π
- $X(t)$: configuration of the voter model at time t
- Then, $M(t) = \langle \pi | X(t) \rangle$ is a bounded martingale
- Optional stopping theorem yields:

$$P(consensus\ value = 1) = \langle \pi | X(0) \rangle$$

Bounds on the coalescence time: continuous time model

- Random walks on n-vertex graph with $q(v, w) = {}^{1}/_{\deg(v)}$

Theorem (Hassin & Peleg, 2001): On any connected graph,

$$E[T] = O(n^3 \log n)$$

- Close to best possible: $\Theta(n^3)$ in general
  - e.g., dumbbell graph: two cliques of $n/4$ vertices, joined by a path of $n/2$ vertices
- Better results available for specific types of graphs

Bound on the coalescence time: discrete time model

- Definitions and notation:
  - λ_2: second eigenvalue of random walk transition probability matrix
  - $\nu = \frac{\sum_{v \in V} \deg(v)^2}{nd^2}$, where $d =$ average degree,

Theorem(Cooper, Elsasser, Ono & Radzik, 2016)

$$E[T] \leq \frac{const.}{1-\lambda_2}\left(log^4 n + \frac{n}{\nu}\right)$$

Corollaries

- Coalescence time is linear in n for expanders
- Coalescence time is sublinear in n for power-law (scale-free) random graphs

Key ideas behind proof

- Fix $k^* = k^*(n)$
- Split coalescence time into time to go from n particles to k^* and then from k^* to 1, and bound each of them separately
- First part: Bound probability that there are k^* particles that avoid meeting each other for a given time
- Second part: Instead of k independent random walks on G, consider one random walk on tensor product graph G^k
- Meeting time of two walks on G is equivalent to hitting time of a specific subset on G^k

Product graph and hitting times

- Vertices of G^k are k-tuples $(v_1, v_2, \dots, v_k)$
- Edge $(\boldsymbol{v}, \boldsymbol{w})$ exists if (v_i, w_i) exists for all i
- $S_k = \left\{\boldsymbol{v} : \ v_i = v_j \ for \ some \ i \neq j\right\}$: set of diagonal vertices
- Γ : graph obtained by contracting S_k to single vertex, γ
- Want to bound hitting time of γ from arbitrary initial condition
- Bound it by sum of mixing time, and hitting time started from stationarity
- Key fact: $\lambda_2(G) = \lambda_2\left(G^k\right) \geq \lambda_2(\Gamma)$

Moran model with selection

- $N(t)$: number of A alleles at time t
- Define $M(t) = (1+s)^{-N(t)}$

- $M(t)$ is a martingale
- Optional Stopping Theorem implies fixation probability is given by

$$P(N(T) = n | N(0) = 1) \to \frac{s}{1+s} \quad \text{as} \quad n \to \infty$$

- Expected time to fixation is $E[T] \approx \log n$

Moran model with selection on a graph

- Very little known in general
- No good tools – neither duality nor martingales

Lieberman, Hauert & Nowak (2005):

- Fixation probability is exactly the same for *regular* graphs/digraphs as for the complete graph
- It is approximately the same if the graph is close to regular, i.e., ratio of maximum to minimum degree is close to 1

Majority consensus

Problem Statement

- Given a population of n voters, with opinions in $\{0,1\}$
- Objective: determine the majority vote
  - quickly, i.e., in $polylog(n)$ time
  - reliably, i.e., with high probability, but not necessarily equal to 1
  - using decentralised algorithms
- Motivation: consensus in large distributed systems

Voters with internal states

- Each voter in one of 3 possible states, 0, 1 or undecided
- Contacts between voters as in classical voter model
- But update rule on opinions is different
  - if a voter contacts someone with the same opinion, or an undecided voter, his opinion stays unchanged
  - if a 0 contacts a 1 (or vice versa), he becomes undecided
  - an undecided voter copies the opinion of whoever he contacts

- *Q: What is the probability of reaching consensus on 0/1?*
- *Q: How long does it take?*

Results for 3-state model

- Definitions and notation:
  - α, $1-\alpha$: initial fraction of voters in state 0, 1
  - p_α: probability of reaching consensus on 1
  - $H(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$: entropy of Bern(α) random variable
  - $D(\alpha; {}^1/_2) = \log 2 - H(\alpha)$: relative entropy of Bern(α) wrt Bern(${}^1/_2$)

Theorem (Perron, Vasudevan and Vojnovic, 2008): If $\alpha > {}^1/_2$, then

$$p_\alpha \approx \exp\left(-nD\left(\alpha; {}^1/_2\right)\right)$$

Expected time to reach consensus is $O(\log n)$

Remarks about 3-state model

- Probability of reaching consensus on minority value is small
  - decays exponentially in population size, rather than being constant, as in classical voter model
- Consensus is reached quickly
  - in time logarithmic in population size, as opposed to linear in population size for the classical voter model
- Algorithm is fully decentralised

Open Questions

- Is there an advantage to having even more internal states?
- What happens if voters aren't identical and only some have internal states?

Alternative model: Polling and majority rule

- No internal states
- Voter activity is as in classical model, but …
- … each voter, on becoming active, polls two other voters, and only changes opinion if both have opposite opinion
- More generally, polls m other voters, and only changes state if at least k have opposite opinion
- Yields a family of models parametrised by (m, k)
- Vector (m, k) can be random, to model individual differences

Results for polling model: $k = m$

- Definitions and notation:
  - α, $1-\alpha$: initial fraction of voters in state 0, 1
  - p_α: probability of reaching consensus on 1

Theorem (Cruise and G., 2014): If $\alpha > {}^1/_2$, then

$$p_\alpha \approx \exp\left(-n(m-1)D(\alpha; {}^1/_2)\right)$$

Expected time to reach consensus is $O(\log n)$ for fixed m, but grows exponentially in m

Results for polling model: $k < m$

- Definitions
  - Z_x: Binomial random variable
  - $g(x) = \frac{xP(Z_x \le m-k)}{(1-x)P(Z_x \ge k)}$

Theorem (Cruise and G., 2014):

$$\frac{1}{n}\log p_\alpha \to \int_{1-\alpha}^{1/2} \log g(x)dx \;\; as\, n \to \infty$$

Time to consensus is $O(\log n)$ for fixed m and k

Results for polling model: random (m, k)

- Define $g(x)$ as before, but now taking expectations over the joint distribution of (m, k)
- Then, theorem still holds, with this modified g

- Implication:
  - Have exponential decay of error probabilities, and fast convergence to majority, even if only a fraction of voters use the majority rule, and the rest behave as in the classical voter model

Main ingredients of proof

- Number of 1 votes follows a biased random walk
- Bias is towards the closer boundary, 0 or n, and gets stronger as you get closer to the boundary
- Want to compute probability of hitting boundary at n before boundary at 0
- Exploit well-known connection between random walks and electrical networks to compute hitting probabilities
- Hitting times bounded in terms of number of visits to each state before hitting the boundary

Two-sampling on general graphs

- Cooper, Elsasser, Radzik, Rivera, Shiraga (2015):
- Time is discrete
- In each time step, each node samples two neighbours independently, with replacement – if both have same opinion, adopts that opinion
- Sampling can be uniform, or weighted by given edge weights:

$$P(u, v) = \frac{w(u, v)}{w(u)}$$

Definitions and notation

- A, B : subsets of vertices with initial opinion 0, 1

$$\lambda = \max\{\lambda_2(P), |\lambda_n(P)|\}$$

$$\nu = \frac{n \sum w^2(v)}{w(G)^2}$$

$$\epsilon_0 = \frac{|w(A) - w(B)|}{w(G)}$$

- δ, ξ : arbitrary small constants; K : arbitrary large constant

Main result

- Theorem (Cooper et al., 2015)
- If $\lambda < \frac{1}{\sqrt{2}} - \delta$ and $\epsilon_0 > 2\lambda^2$ and $\epsilon_0^2 > \frac{\nu K \log n}{n}$, then whp voting is completed in $O(\log n)$ rounds, and the winner is the opinion with larger initial weight.
- If $\lambda = o(1)$ and $\lambda^\xi > \frac{\nu K \log n}{n}$, then whp voting is completed in $O(\log \frac{1}{\epsilon_0}) + O(\log \log \frac{1}{\lambda}) + O(log_{1/\lambda} n)$ rounds and the winner is the opinion with larger initial weight