Seminar 5 – Population Protocols: General Computation and Leader Election Distributed Computing in Anonymous Dynamic Systems

Giovanni Viglietta

Rome – March 12, 2024

Syllabus

- Anonymous Networks
 - Introduction and basic algorithms for static networks
 - Dynamicity and history trees
 - Optimal computation in networks with and without leaders
 - Computation in dynamic congested networks
- Population Protocols
 - Introduction and basic algorithmic techniques
 - Leader election in Mediated Population Protocols
- Mobile Robots
 - Gathering and Pattern Formation in the plane
 - Meeting in a polygon by oblivious robots

Exam

Pre-recorded 10-minute presentation video on one of the papers that will be suggested at the end of the course.

- Introduction to Population Protocols
- Computable predicates
- Mediated Population Protocols
- One-way and faulty models



Setting: a set of finite-state agents.



Pairs of agents interact in a non-deterministic order...















- A **Population Protocol** is a network of anonymous agents where:
 - Communications are sequential (i.e., only one pair of agents interacts at any time),
 - Interactions are asymmetric (i.e., there is an initiator and a responder).
 - Each agent has a constant amount of internal memory.

So, if internal states belong to a finite set S, a *transition function* has the form $f: S \times S \rightarrow S \times S$.

There are two notions of *fairness* of interactions:

- Local fairness: Any pair of agents interacts infinitely often (in both ways).
- Global fairness: If a configuration of *all agents* is potentially reachable for infinitely many times, it is eventually reached.

Population Protocols: Leader Election

Example. Leader Election protocol rules:

 $(L, L) \mapsto (L, N)$ $(L, N) \mapsto (L, N)$ $(N, L) \mapsto (N, L)$ $(N, N) \mapsto (N, N)$

If all agents start in state L, eventually there will be only one in state L (the *leader*). This protocol works under local fairness.

Population Protocols: Majority

Example. Majority protocol. Agents start in state "red" or "blue". Eventually, they reach a configuration where all agents are in state "yes" if # red > # blue, and in state "no" otherwise. Rules: (red, blue) \mapsto (no, no) (eliminates all "red"s or all "blue"s) (red, no) \mapsto (red, yes) ("red" changes answer to "yes") (blue, yes) \mapsto (blue, no) ("blue" changes answer to "no") (yes, no) \mapsto (no, no) (takes care of ties)

This protocol requires global fairness (why?).

Population Protocols: Computable Predicates

Assuming that the initial states are integers (modulo m), what predicates can be computed by Population Protocols?

A *predicate* is a function whose output is "yes" or "no", and all agents must eventually return the correct output based on the multiset of initial states.

Theorem (Angluin et al., PODC 2006)

The only predicates computable by Population Protocols are:

- $\sum_{i} c_i x_i \ge a$, where a, c_i 's are integer constants (generalization of Majority),
- $\sum_{i} c_i x_i \equiv a \pmod{b}$, where a, b, c_i 's are integer constants,
- Boolean combinations of the above predicates (\neg, \lor, \land) .

These can also be characterized in terms of *Presburger arithmetic*: the predicates in first-order logic using the symbols +, 0, 1, \neg , \lor , \land , \forall , \exists , =, <, (,), plus variables.



Assume that agents can only interact through specific links.



Pairs of adjacent agents interact in a non-deterministic order...











Mediated agents: we add ports with (finite) states.



Mediated agents: we add ports with (finite) states.



The transition function affects both agent and port states.

f(A, B, a, b) = (C, D, c, d)

The transition function affects both agent and port states.



Each agent has a port for each neighbor.

We distinguish two types of scheduler:

- **Recurrent:** each pair of neighboring agents interacts infinitely often (in both directions)
- *k*-**Bounded:** it is recurrent and, between two consecutive interactions of the same pair of agents, no other pair interacts more than *k* times

Note: under a <u>1-bounded</u> scheduler, the sequence of interactions is periodic

A protocol can be:

- Stable: eventually, no agent changes state
- **Terminating:** eventually, all agents are in a *terminal state* (i.e., "explicit stability")

Usually, with the <u>recurrent</u> scheduler protocols are <u>stable</u>; with the <u>k-bounded</u> schedulers, they are terminating

Leader Election Problem: all agents start in the same state, and eventually there is a unique agent in a *leader state*

- Complete graphs
- Complete bipartite graphs
- Trees

Applications of a Unique Leader:

- Token circulation
- Construction of a shortest-path spanning tree
- Stability detection (turning stable protocols into terminating ones)
- Equivalence of k-bounded schedulers for all k > 1



Theorem: in the complete graph K_n , it is possible to elect a leader under the <u>recurrent</u> scheduler.



Initially, all agents have the leader state.



When two leaders interact, one is "eliminated".



When two leaders interact, one is "eliminated".


When two leaders interact, one is "eliminated".



When two leaders interact, one is "eliminated".



When two leaders interact, one is "eliminated".



Otherwise, nothing happens.



Otherwise, nothing happens.



Otherwise, nothing happens.



Eventually, only one leader remains.



Theorem: in $K_{m,n}$, it is possible to elect a leader under the <u>1-bounded</u> scheduler if and only if m and n are coprime.



Suppose that m and n are coprime.













The next time it encounters the marked edge, it knows that a new period has started.



In the first phase, we construct a maximal matching.















When an unmatched agent sees only matched agents for an entire period, it knows that the matching is maximal.



These unmatched agents assume a "reset" state.



After another period, whoever sees a "reset" agent becomes unmatched again.



After another period, whoever sees a "reset" agent becomes unmatched again.



In the next period, the agents that are still matched become "eliminated".















Since m and n are coprime, eventually only one agent will remain available.



When this agent sees that all its neighbors are "eliminated", it becomes the leader. This protocol is terminating.



Suppose that m and n have a common divisor d > 1.


We partition each side of the network into groups of d agents, and we assign all agents the same initial state.



The scheduler chooses two groups on opposite sides, and activates the agents according to a perfect matching.



Then it chooses another perfect matching, and so on, until all pairs of neighbors have been activated.



Then it chooses another perfect matching, and so on, until all pairs of neighbors have been activated.



Then it does the same with two other groups, and so on, until all pairs of neighbors have been activated.



Every d interactions, all agents in the same group have the same state. Hence a leader cannot be elected.



Theorem: in a tree, it is possible to elect a leader under the <u>recurrent</u> scheduler.



When two agents meet, they form a small rooted tree, where the root is a leader.



Arrows are encoded as port states.



New agents may join existing trees, and a forest is formed.



When two trees merge, one agent becomes "busy". Its task is to tell its leader that it is no longer a leader.



The parent of a busy agent becomes busy too, and reverses the corresponding arrow.



The child then ceases to be busy.



A busy agent rejects all requests to merge.



A busy agent rejects all requests to merge.











Agents that are no longer busy accept new merge requests.



Eventually, only one leader is left, and the whole tree is oriented toward it.



If the scheduler is \underline{k} -bounded, a leaf eventually knows that it is a leaf. A non-leader leaf can safely terminate.



If the scheduler is \underline{k} -bounded, a leaf eventually knows that it is a leaf. A non-leader leaf can safely terminate.



If an agent's children all have terminated, the agent eventually realizes and terminates.



Eventually, all non-leader agents terminate, and hence the protocol is terminating.



Suppose there is a unique leader and we want to make it "visit" the entire network.



By that we mean that the leadership is "transferred" to a different agent during an interaction.



When an agent that has never been leader meets the leader, it takes the leadership.



As the leader goes, it leaves a "trail" of arrows.



When a new agent meets an agent that has already been leader, it becomes a "summoner".



When a new agent meets an agent that has already been leader, it becomes a "summoner".



The parent of a summoner becomes a summoner as well, and reverses the corresponding arrow.



The parent of a summoner becomes a summoner as well, and reverses the corresponding arrow.



When the leader meets a summoner, it gives it the leadership.



A summoner ignores all requests from new agents.



The leader keeps following the arrows through summoners, reversing them as it goes.


Different agents may summon the leader in parallel, but they never interfere with each other.



Different agents may summon the leader in parallel, but they never interfere with each other.



Different agents may summon the leader in parallel, but they never interfere with each other.











Eventually, the leader visits the entire network. As a byproduct, a rooted spanning tree has been constructed.



Note that this spanning tree may not be balanced.



Say we want to construct a better spanning tree rooted at the leader, under the \underline{k} -bounded scheduler.



When a new agent interacts with the leader, it becomes a "leaf".



When a new agent interacts with the leader, it becomes a "leaf".



When a new agent interacts with the leader, it becomes a "leaf".



Since the scheduler is \underline{k} -bounded, the leader knows when all its neighbors are leaves.



The leader issues a "new level" command along the tree.











When they realize that all their neighbors have been included in the spanning tree, they send a "done" message to the leader.



When they realize that all their neighbors have been included in the spanning tree, they send a "done" message to the leader.



The leader issues another "new level" command, which is forwarded along the spanning tree.



The leader issues another "new level" command, which is forwarded along the spanning tree.



A new level of the spanning tree is constructed.



The leaves that are unable to expand assume a terminal state and send a "terminated" message toward the leader.



The leaves that are unable to expand assume a terminal state and send a "terminated" message toward the leader.



The leaves that are unable to expand assume a terminal state and send a "terminated" message toward the leader.













When the leader receives a "terminated" message from all its children, it terminates.

Application: Stability Detection



Say we have a leader and a spanning tree, and we want to detect (under the <u>k-bounded</u> scheduler) when a protocol P stabilizes.

Application: Stability Detection



Whenever a new edge is activated, its endpoints "simulate" a transition according to P.

Application: Stability Detection



Whenever a new edge is activated, its endpoints "simulate" a transition according to P.


Whenever a new edge is activated, its endpoints "simulate" a transition according to P.



These edges are marked, so the corresponding simulated interaction does not occur twice.



These edges are marked, so the corresponding simulated interaction does not occur twice.



If a simulated interaction over an edge leaves the simulated states unchanged, the edge is marked as "stable".



If a simulated interaction over an edge leaves the simulated states unchanged, the edge is marked as "stable".



Since the scheduler is \underline{k} -bounded, an agent eventually realizes that it has interacted with all its neighbors.



When this happens, the agent becomes "done".



If all the children of a "done" agent (in the spanning tree) are "done", the agent forwards a "done" message to its parent.



If all the children of a "done" agent (in the spanning tree) are "done", the agent forwards a "done" message to its parent.



If all the children of a "done" agent (in the spanning tree) are "done", the agent forwards a "done" message to its parent.



Eventually, the leader receives "done" messages from all its children.



At this point, the leader broadcasts a "reset" message.



At this point, the leader broadcasts a "reset" message.



All edges that are incident to a "reset" agent become unmarked.



The "reset" message is forwarded along the spanning tree.



Eventually, the whole network is reset. The leader is notified, and starts a new simulation phase.



If P is <u>stable</u>, eventually all edges will be marked as "stable".



If P is <u>stable</u>, eventually all edges will be marked as "stable".



All agents send "done" and "stable" messages to their parents.



Eventually, the leader receives "done" and "stable" messages from all its children.



The leader then terminates and broadcasts a "terminate" message, which is forwarded along the spanning tree.



The leader then terminates and broadcasts a "terminate" message, which is forwarded along the spanning tree.



This converts the stable protocol P into a terminating one.



Note: The simulated schedule activates all edges of the network in some order, then it activates them again in some other order, etc.



So, between two activations of an edge (say, a), each other edge is activated at most twice.



So, between two activations of an edge (say, a), each other edge is activated at most twice.



So, between two activations of an edge (say, a), each other edge is activated at most twice.



It follows that the simulated schedule is $\underline{2\text{-bounded}}$.



So, the protocols that work under the <u>2-bounded</u> scheduler also work under all <u>k-bounded</u> schedulers, for all k > 2.



Theorem: in every network where a leader can be elected, the <u>k-bounded</u> schedulers are all equivalent, for k > 1.

One-Way and Faulty Models



The traditional interaction model is called **Two-Way**.



Immediate Observation: only the second agent transitions.



Immediate Transmission: the first agent detects proximity.



I1: IT with omission faults, no detection.



I2: IT with omission faults, proximity detection.


 I_3 : IT with omission faults, reactor-side omission detection.



I_4 : IT with omission faults, starter-side omission detection.



 T_1 : TW with omission faults, no detection.



T₂: TW with omission faults, starter-side omission detection.



T₃: TW with omission faults, omission detection by both sides.



Theorem: all possible models obtained by combining one-way and two-way interactions with omission detection and proximity detection, starter-side or reactor-side, fall into one of these classes.



We seek to *simulate* two-way interactions in weaker models.





The simulating agents have a simulated state and a work state.





Typically, an interaction determines a change in the work state.



Typically, an interaction determines a change in the work state.





Typically, an interaction determines a change in the work state.





Occasionally, changes in the simulated state may occur.





Occasionally, changes in the simulated state may occur.





Occasionally, changes in the simulated state may occur.





These have to mimic transitions in the simulated TW protocol.



Globally, we want to pair up simulated states transitions...



... in a way that is compatible with the simulated TW protocol.



Suppose the simulating agents have **infinite memory**: what models can simulate *all* TW population protocols?



In **IT**, we can implement a *token-passing* technique that can be used to simulate two-way interactions.



In **T**₃, it is impossible to simulate a two-way protocol for the *pairing problem*.



As a consequence, simulation is impossible also in the weaker interaction models.

Results: unique IDs



Suppose the simulating agents have *unique IDs* as part of their initial state.

Results: unique IDs



In **IO**, we can implement a *locking mechanism*, along with a *rollback process* to avoid deadlocks.

Results: unique IDs



As a consequence, simulation is possible also in the stronger interaction models.

Results: knowledge of n



Suppose the simulating agents know the size of the system, n, and have $O(\log n)$ bits of internal memory.

Results: knowledge of n



In **IO**, we can implement a *naming algorithm* that eventually gives each agent a unique ID.

Results: knowledge of n



When an agent has ID n, the system starts executing the previous unique-ID simulation protocol.



Suppose that the simulating agents are given an upper bound b on the number of faulty interactions in the system.



In I_3 and I_4 , we can extend the token-passing technique by splitting each token into b + 1 parts.



In T_1 , I_1 , and I_2 , it is impossible to simulate the pairing protocol, even for b = 1.



Open problem: is it possible to simulate all TW protocols in T_2 , given an upper bound on the number of faulty interactions?