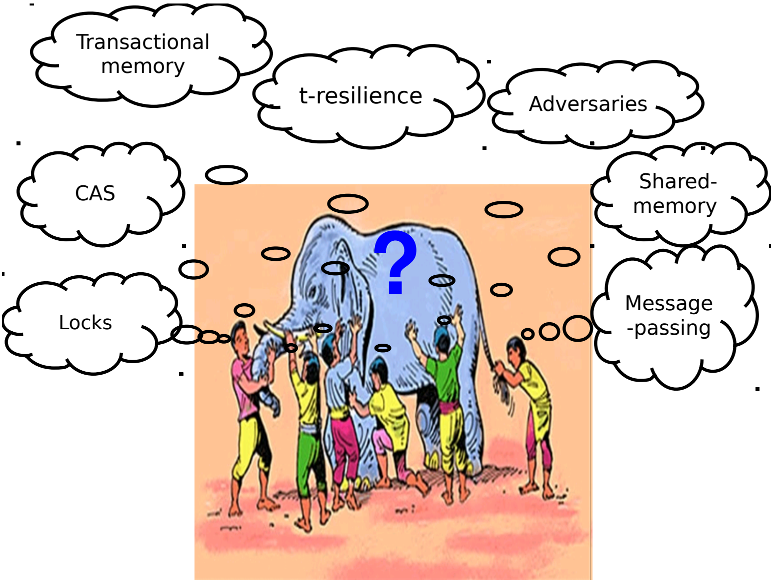
Affine tasks for distributed computing models

Petr Kuznetsov[1] and Thibault Rieutord[1]
Joint work with Eli Gafni[2] and Yuan He[2]

[1] LTCI, Télécom ParisTech, Université Paris Saclay
[2] UCLA

8th of March 2017

A large diversity of models



Set of runs

Processes steps interleaving

Define rules on processes steps ordering:

- wait-free model
- t-resilience
- Adversaries
- ...

k-concurrency

Between the first and last step of a process, at most $k - 1$ other processes performed steps.

Shared memory + distributed objects

Diversity of available objects

- Test-and-Set
- Stacks
- Compare-and-Swap
- ...

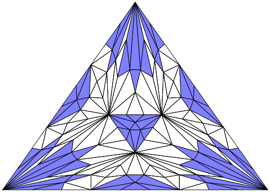
k-set-consensus
Processes propose a value v, and, if correct, returns a decision, such that, a decision is a proposed value and at most k distinct values are returned.

Model as (affine) tasks

Takeaway

A (long-lived, non-compact) model can be matched by a (one-shot, compact) task.

***Fair* Adversaries have a matching affine task.**



Superseeds affine tasks for t-resilience [SHG16] and k-concurrency [GHKR16].

Overview of the Presentation

1. Agreement functions
2. Topological Representations and the IIS Model
3. Affine Tasks for Fair Adversaries
4. Sketch Proof of the Equivalence

Agreement function

Agreement function:

The *agreement function* of a model M is a function $\alpha : 2^{\Pi} \rightarrow \{0, \ldots, n\}$, such that for each $P \in 2^{\Pi}$, in the set of runs of M in which no process in $\Pi \setminus P$ participates, iterations of $\alpha(P)$-set consensus can be solved, but $(\alpha(P) - 1)$-set consensus cannot.
By convention it is equal to 0, if no (infinite) runs with participating set P exists in M.

Monotonicity: For any model, if $P \subseteq P'$ then $\alpha(P) \leq \alpha(P')$.

Generic agreement function model

α-model:

The α*-model* is the set of runs in which, the participating set P is such that:

- $alpha(P) \geq 1$;
- at most $\alpha(P) - 1$ participating processes are faulty.

Universality of the α-model

[KR17]

Monotonic α-model belong to the *weakest class* of models with agreement function α.

Proof sketch:

- The agreement function of the α-model is α;
- A model with agreement function α can solve an α-adaptive set consensus.
- Any task solvable in the α-model can be solved in a model M with agreement function α.

Fair Adversaries

Adversaries[DFGT09]:

- An adversary $\mathcal{A}$ is a set of processes sets, called live-sets.
- An $\mathcal{A}$-compliant run is an infinite run where the set of correct processes is a live set of A.
- The adversarial $\mathcal{A}$-model is the set of $\mathcal{A}$-compliant runs.
- The agreement function of an adversary is $\alpha_{\mathcal{A}} = \mathit{setcon}(\mathcal{A}|_P)$ [GK10].

Fair Adversaries

Adversaries[DFGT09]:

- An adversary $\mathcal{A}$ is a set of processes sets, called live-sets.
- An $\mathcal{A}$-compliant run is an infinite run where the set of correct processes is a live set of A.
- The adversarial $\mathcal{A}$-model is the set of $\mathcal{A}$-compliant runs.
- The agreement function of an adversary is $\alpha_{\mathcal{A}} = \mathit{setcon}(\mathcal{A}|_P)$ [GK10].

Fair adveraries:

An adversary $\mathcal{A}$ is fair if and ony if:
$\forall P, Q \subseteq \Pi, \mathit{setcon}(\{L \in \mathcal{A}|_P, L \cap Q \neq \emptyset\}) = \mathit{min}(\mathit{setcon}(\mathcal{A}|_P), |Q|)$.

Equivalence with α-model[KR17]

Symmetric and superset-closed adversaries are fair adversaries.

A fair adversary with agreement function α is equivalent to the α-model.

The agreement function α of an adversary is *regular*:

$$\forall P, Q \subseteq \Pi, P \cup Q = \emptyset, \alpha(P \cup Q) \leq \alpha(P) + |Q|.$$

Immediate snapshot object

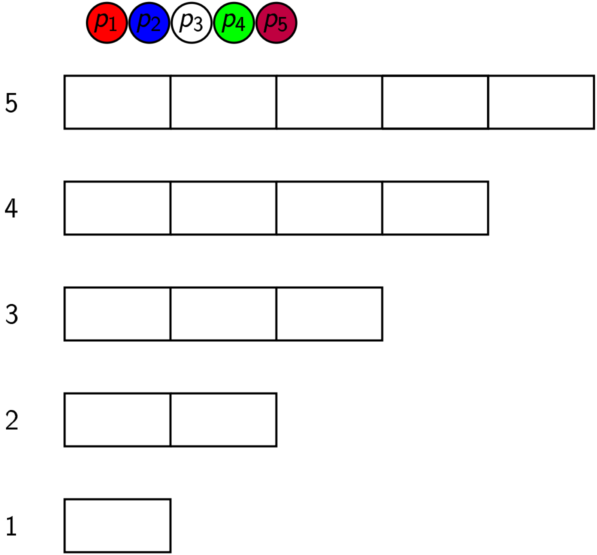
An object with a single operation:

- Takes a value v_i;
- Returns a set of submitted values V_{ir}.

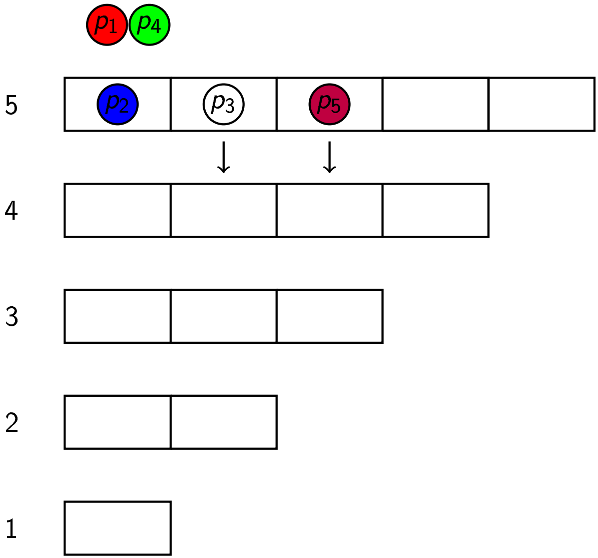
Immediate Snapshot Properties:

- *self-inclusion:* $v_i \in V_{ir}$;
- *containment:* $(V_{ir} \subseteq V_{jr}) \vee (V_{jr} \subseteq V_{ir})$;
- *immediacy:* $v_i \in V_{jr} \Rightarrow V_{ir} \subseteq V_{jr}$.

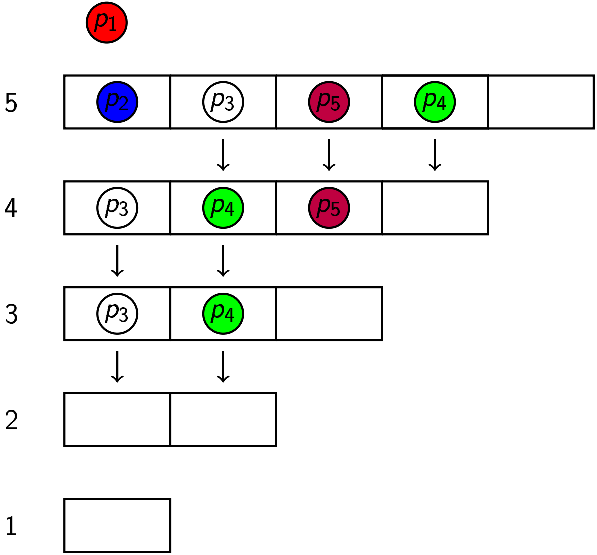
Immediate snapshot algorithm



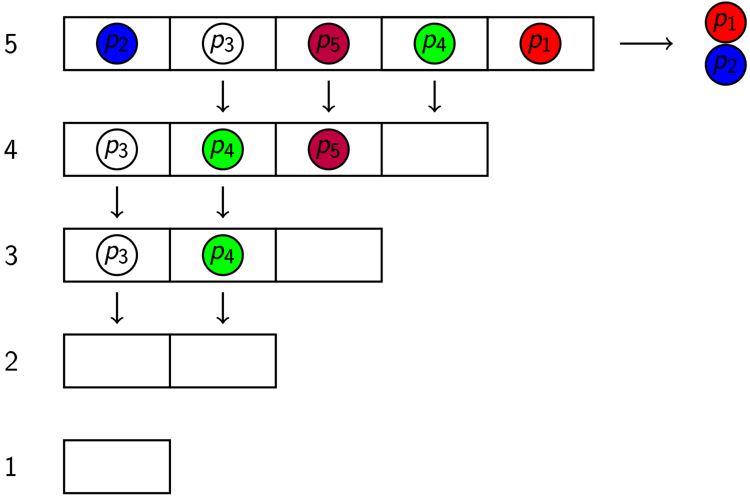
Immediate snapshot algorithm



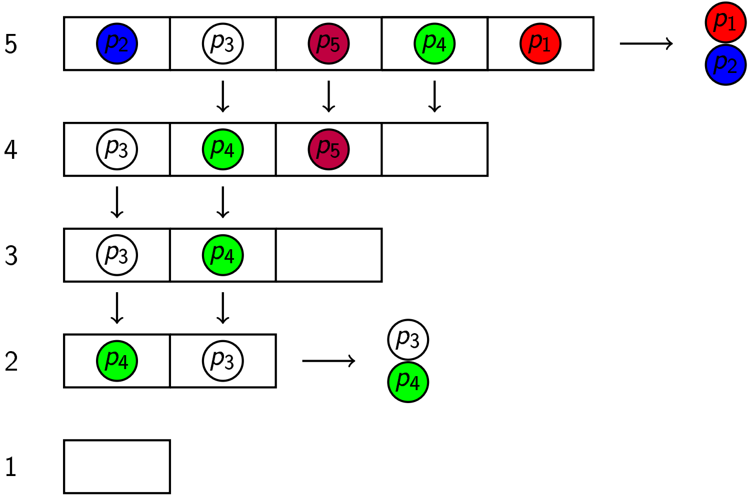
Immediate snapshot algorithm



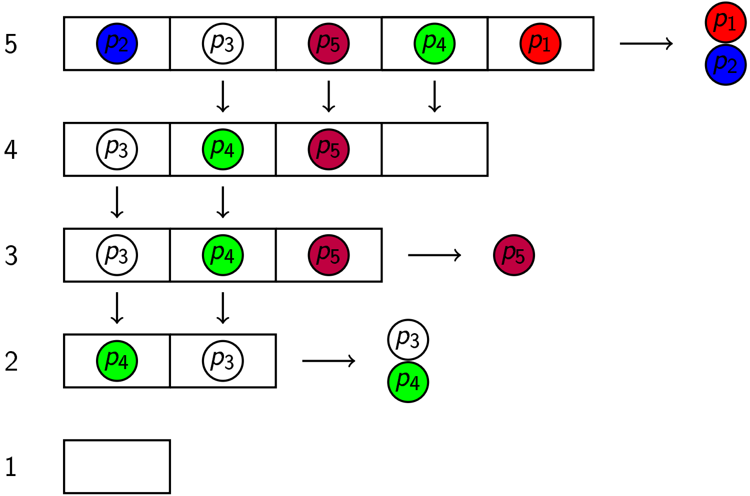
Immediate snapshot algorithm



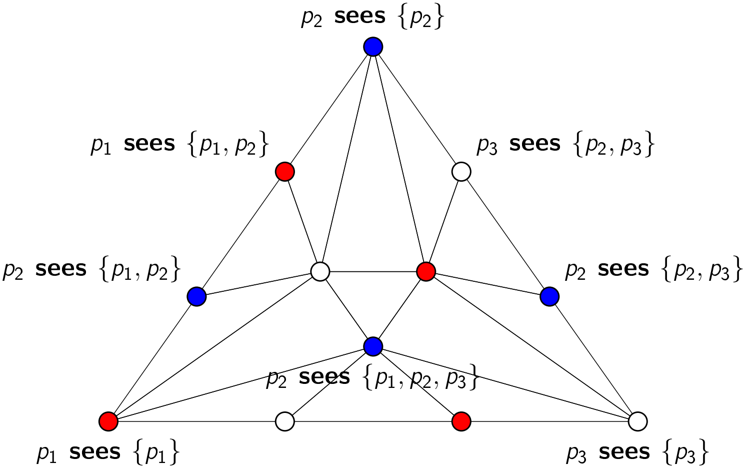
Immediate snapshot algorithm



Immediate snapshot algorithm

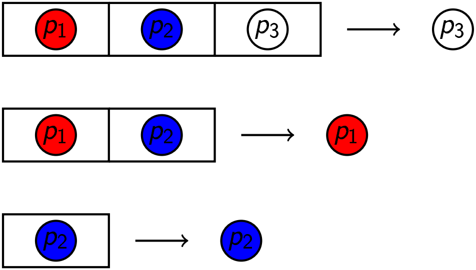


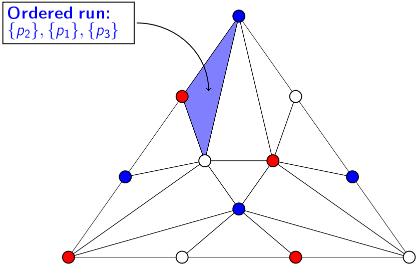
Topological representation



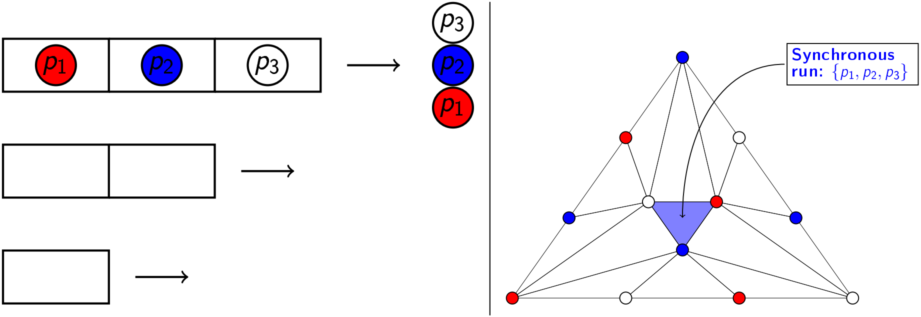
Standard chromatic subdivision

Example of IS runs





Example of IS runs



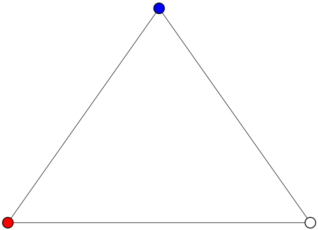
Tasks

Distributed task $(\mathcal{I}, \mathcal{O}, \Delta)$:

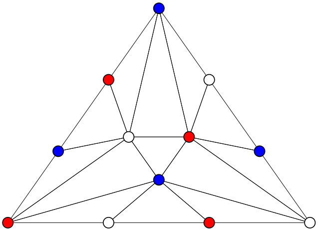
- $\mathcal{I}$: Input complex, i.e., set of valid inputs combinations;
- $\mathcal{O}$: Output complex, i.e., set of valid outputs combinations;
- Δ: Carrier map, function from $\mathcal{I}$ to $2^{\mathcal{O}}$:
  Subset of outputs valid for an input combination.

Simplex agreement task

$\mathcal{I}$:

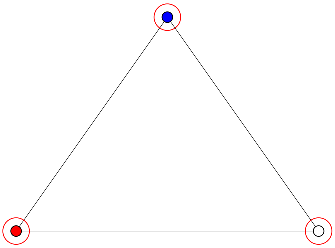


$\mathcal{O}$:

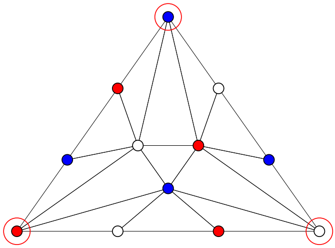


Simplex agreement task

$\mathcal{I}$:

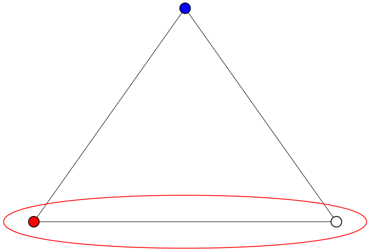


$\mathcal{O}$:

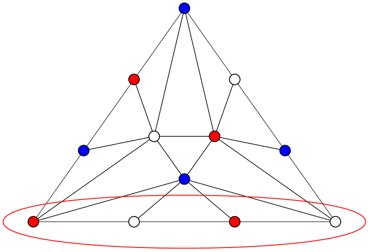


Simplex agreement task

$\mathcal{I}$:



$\mathcal{O}$:

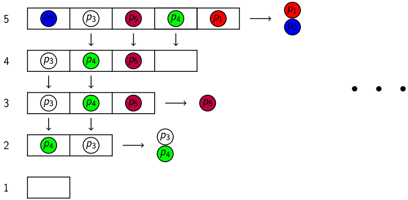


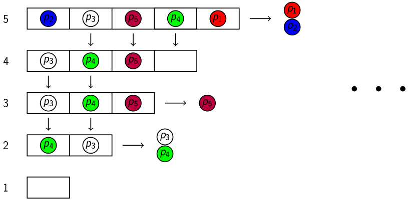
IIS as iteration of the task

Going through a sequence of immediate snapshots

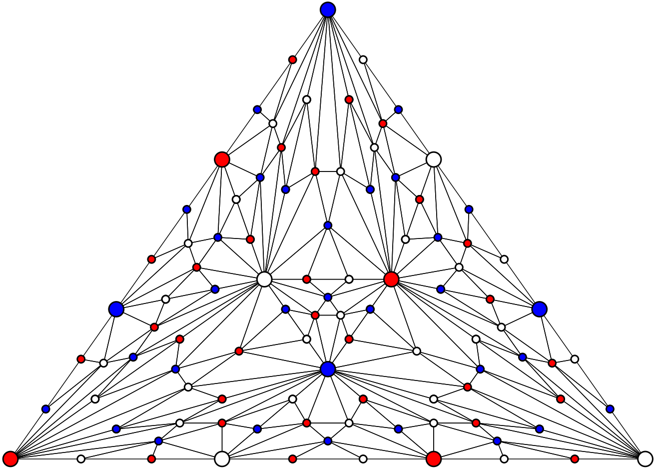
(IIS) consists in an infinite sequence of memories that can each be accessed only once by any process.

$$IS^1, IS^2, \ldots, IS^m, \ldots$$





Iterated subdivisions



2^{nd} **Iteration of the standard chromatic subdivision**

Wait-free task computability

Read-write (RW) model and IIS are equivalent[BG93,BG97,GR10]

A task is solvable in IIS if and only if it is wait-free solvable in RW.

Asynchronous computability theorem[HS93]

A task $(\mathcal{I}, \mathcal{O}, \Delta)$ is wait-free read-write solvable if and only if there is a chromatic simplicial map from a subdivision $\chi(\mathcal{I})$ to $\mathcal{O}$ carried by Δ.

Affine tasks

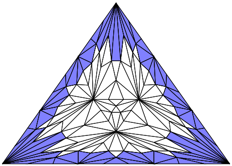
IS is the matching task for wait-free runs

What about model stronger than wait-free?

Use restrictions of the wait-free task

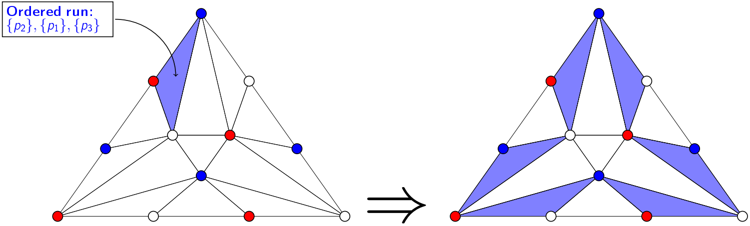
A task defined as the simplex agreement task:

- $\mathcal{I}$: n-dimensional simplex s.
- $\mathcal{O}$: $L \subset Chr^m(s)$.
- Δ: $\Delta(\sigma) = \{\sigma' \in \mathcal{O}, |\sigma'| \subseteq |\sigma|\}$.



Multiple Iterations (may be) required

1-OF adversary cannot be captured as an affine task of $Chr(s)$:



The corners remain connected under iterations.

Contention Simplices.

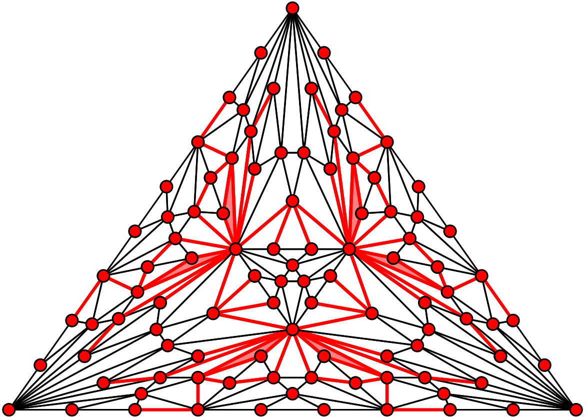
2-Contention simplex:

$\sigma \in Chr^2\mathbf{s}$ such that: $\forall v, v' \in \sigma, v \neq v'$:

$$((View^1(v) \subsetneq View^1(v')) \wedge (View^2(v') \subsetneq View^2(v))) \vee$$
$$((View^1(v') \subsetneq View^1(v)) \wedge (View^2(v) \subsetneq View^2(v'))).$$

Set of vertices with a reverse inclusion ordering in the two levels of subdivision.

Contention Simplices for 3 Processes



Critical Simplices.

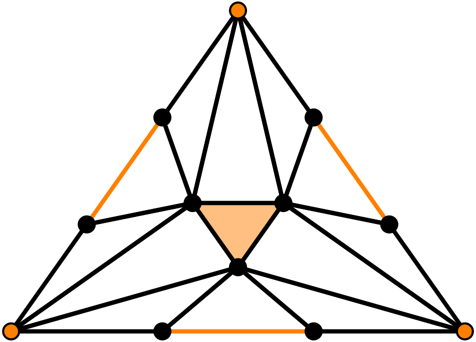
Critical simplex:

$\sigma \in Chr\mathbf{s}$ such that:

$$(\forall v \in \sigma : carrier(v, \mathbf{s}) = carrier(\sigma, \mathbf{s})) \wedge$$
$$(\alpha(\chi(carrier(\sigma, \mathbf{s})) \setminus \chi(\sigma)) < \alpha(\chi(carrier(\sigma, \mathbf{s})))).$$

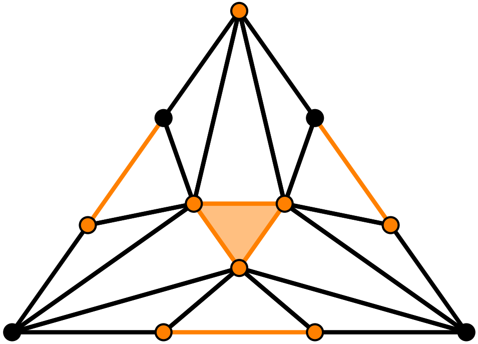
Set of vertices of the first subdivision acting as a "joint" leader.

Critical Simplices Example



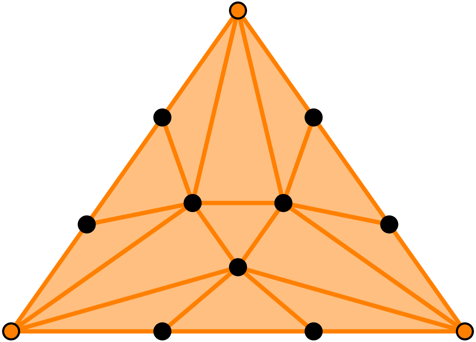
Critical simplices for the 1-Obstruction free adversary.

Critical Simplices Example



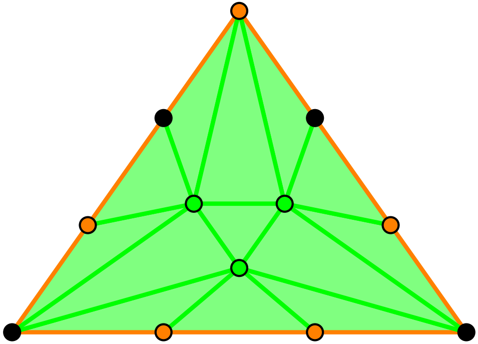
Critical simplices for the superset-closed adversary induced by $\{p_1\}$ and $\{p_2, p_3\}$.

Induced Concurrency Map



Concurrency map for the 1-Obstruction free adversary.

Induced Concurrency Map



Concurrency map for the superset-closed adversary induced by $\{p_1\}$ and $\{p_2, p_3\}$.

Affine task for regular α-models.

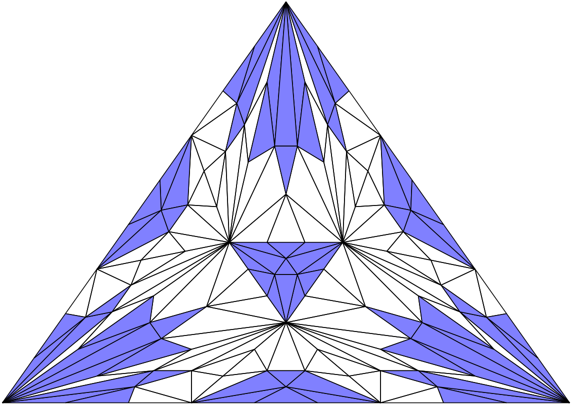
Affine task $\mathcal{R}_\alpha$:

$\sigma \in Chr^2\mathbf{s}, dim(\sigma) = n-1 : \sigma \in \mathcal{R}_\alpha$ if and only if:

$$\forall \theta \subseteq \sigma, \theta' = carrier(\theta, Chr\mathbf{s}) :$$
$$(\theta \in Cont_2) \wedge (\chi(\theta) \cap (\chi(\mathcal{CSM}_\alpha(carrier(\sigma, Chr\mathbf{s}))) \cup \chi(\mathcal{CSV}_\alpha(\theta'))) = \emptyset$$
$$\implies dim(\theta) - 1 \leq Conc_\alpha(\theta').$$

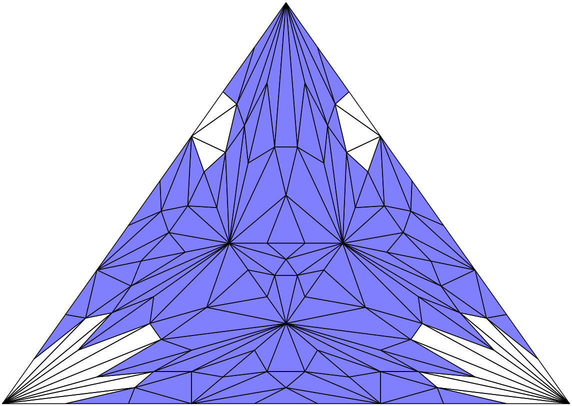
- With $\mathcal{CSM}_\alpha$ the set of critical simplices "members".
- With $\mathcal{CSV}_\alpha$ the "view" associated to a critical simplex.

Example of $\mathcal{R}_\alpha$ (1/2)



Affine task for the 1-Obstruction free adversary.

Example of $\mathcal{R}_\alpha$ (2/2)



Affine task for the superset-closed adversary induced by $\{p_1\}$ and $\{p_2, p_3\}$.

Solving $\mathcal{R}_\alpha$

A simple algorithm:

1. Execute First Immedate Snapshot Algorithm;
2. Write Result to Shared Memory;
3. Wait Until **Condition** is Satisfied;
4. Execute Second Immedate Snapshot Algorithm;
5. Write Result to Shared Memory;

Wait Condition

The wait condition is satisfied when eiter:

- Outputs of IS1 indicate that p is a critical simplex member.
- The following conditions are all satisfied:
  - All members of a critical simplex "associated to" a concurrency level of k. have written an IS2 output.
  - The number of processes without an IS2 output in p IS1 view which are not a critical simplex member is stricly smaller than k.

Validity of the algorithms.

The safety property directly derives from the fact that the wait condition is "stricter" than the required properties:

Affine task $\mathcal{R}_\alpha$:

$\sigma \in Chr^2\mathbf{s}, dim(\sigma) = n-1 : \sigma \in \mathcal{R}_\alpha$ if and only if:

$$\forall \theta \subseteq \sigma, \theta' = carrier(\theta, Chr\mathbf{s}) :$$
$$(\theta \in Cont_2) \wedge (\chi(\theta) \cap (\chi(\mathcal{CSM}_\alpha(carrier(\sigma, Chr\mathbf{s}))) \cup \chi(\mathcal{CSV}_\alpha(\theta'))) = \emptyset$$
$$\implies dim(\theta) - 1 \leq Conc_\alpha(\theta').$$

Algorithm Liveness

Intuition of the liveness validity:

- A process failure can block in IS1 a limited number of critical simplexes.
  (The minimal hitting set size of critical simplices is greater than the resilience level.)
- A process failure cannot block critical simplex members are non-critical simplex number at the same time.

$\Rightarrow$ The number of correct processes with a smaller IS1 view "scales" with the concurrency provided with terminated critical simplexes.

From $\mathcal{R}_\alpha$ to the α-Model

Shared memory simulation:

- Shared memory is simulated using the algorithm from [GR10] using iterated snapshots.
- It is executed on $\mathcal{R}_\alpha$ outputs views, all processes observed directly or indirectly in IS2.
- If a process does not know what to write, it re-write the last written value.
- Processes stop participating when they have obtain a task output.

The simulation [GR10], ensures that eventually, all processes with the smallest round snapshot complete a new memory operation.

From $\mathcal{R}_\alpha$ to the α-Model

α-adaptive set-consensus among active processes A:

At every round, processes execute this algorithm:

- Share every agreement operation current decision estimate to $\mathcal{R}_\alpha$;
- If there is a process in μ_A with a proposal then :
  - Adopt the minimal proposal in μ_A;
- If every process in $\mu_A \cap A$ has a set consensus proposal for a given agreement, then:
  - Write the initial state of every process in μ_A to the shared memory;
  - return the minimal proposal in μ_A;

Simplicial map μ_A:

$$\mu_A(v) = \textbf{ if } (\chi(\mathcal{CSV}_\alpha(carrier(v, Chr\,\mathsf{s}))) \cap A \neq \emptyset)$$

$$\textbf{then } \chi(min(\{carrier(\sigma', \mathsf{s}) : (\sigma' \in CS_\alpha(carrier(v, Chr\,\mathsf{s})) : \chi(carrier(\sigma', \mathsf{s})) \cap A \neq \emptyset)\})$$

$$\textbf{else } \chi(min(\{carrier(v', \mathsf{s}) : (v' \in carrier(v, Chr\,\mathsf{s})) \wedge (dim(v') = 1) \wedge (carrier(v', \mathsf{s}) \cap A \neq \emptyset)\}).$$

Topological Characterization of Task Solvability

Regular α-Model ACT

A task $(\mathcal{I}, \mathcal{O}, \Delta)$ is solvable in a regular α-model if and only if there exists $N \in \mathbb{N}$ and a chromatic simplicial map ϕ that maps from $\mathcal{R}_{\alpha}^{N}(\mathcal{I})$ to $\mathcal{O}$ and is carried by Δ.

Fair Adversaries ACT

A task $(\mathcal{I}, \mathcal{O}, \Delta)$ is solvable in an adversarial $\mathcal{A}$-model if and only if there exists $N \in \mathbb{N}$ and a chromatic simplicial map ϕ that maps from $\mathcal{R}_{\alpha_{\mathcal{A}}}^{N}(\mathcal{I})$ to $\mathcal{O}$ and is carried by Δ.

Conclusion

Compact representation of non-compact models:

- k-concurrency and k-set-agreement[GHKR16];
- t-resilience[SHG16];
- Fair adversaries;
- Regular α-models;
- General Adversaries?
- Collection of k-set-consensus?

Conjecture: possible for all "natural models".

3-process, R/W wait-free solvability of tasks are undecidable[GK95,HR97]

Conjecture: relations between models (affine tasks) are decidable.

Questions?

Thank You!