Stochastic Games

Krishnendu Chatterjee

IST AUSTRIA
Institute of Science and Technology

ARiSE Winter School, Vienna,
Feb, 2012

ARiSE
Austrian Rigorous Systems Engineering

Stochastic Games

- Two-player perfect-information games on finite graphs with randomness in transitions.

- Various sub-classes
  - Brief discussion of applications.
  - Solution techniques.

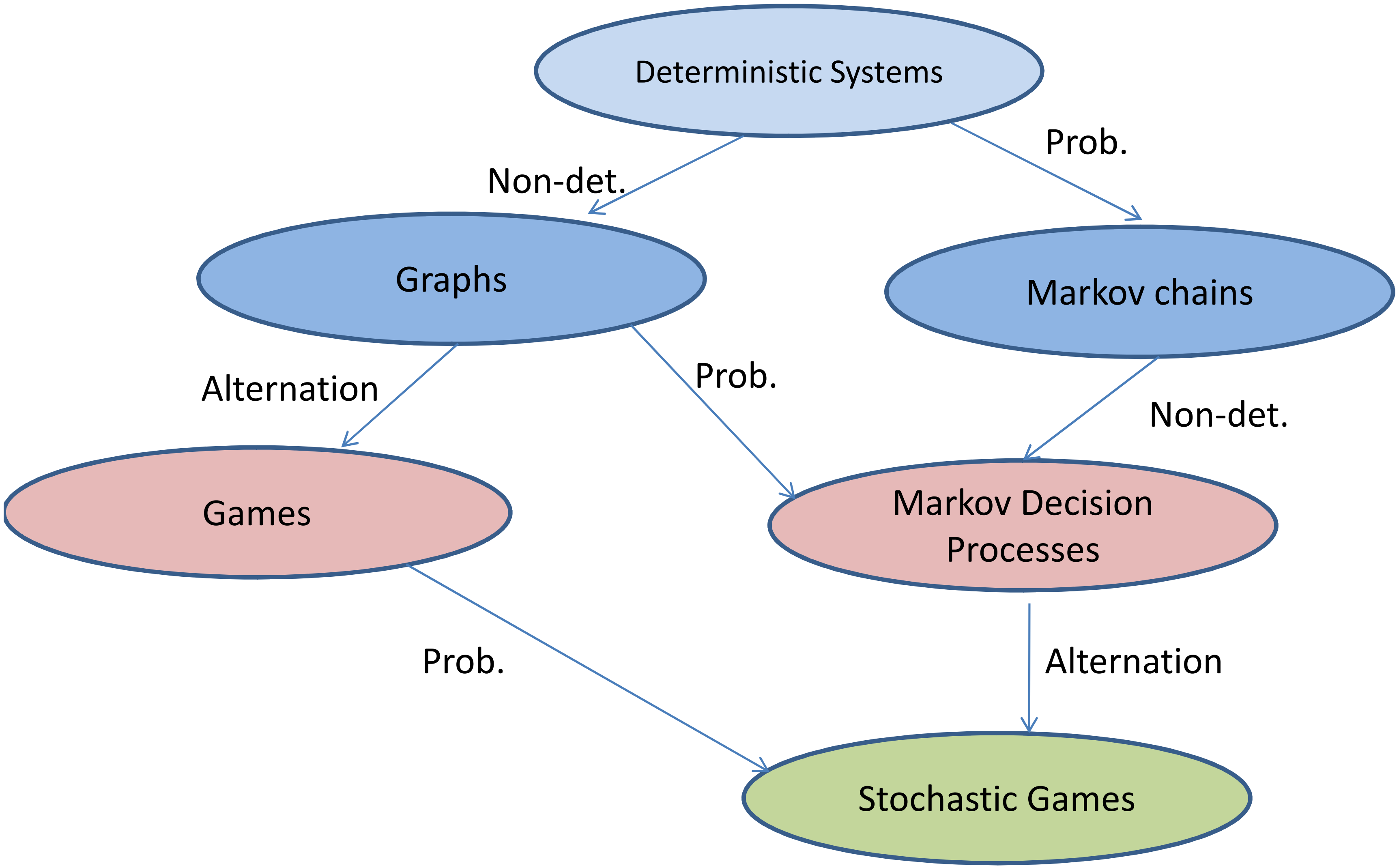
System Analysis

- Formal analysis of systems to prove correctness with respect to properties.

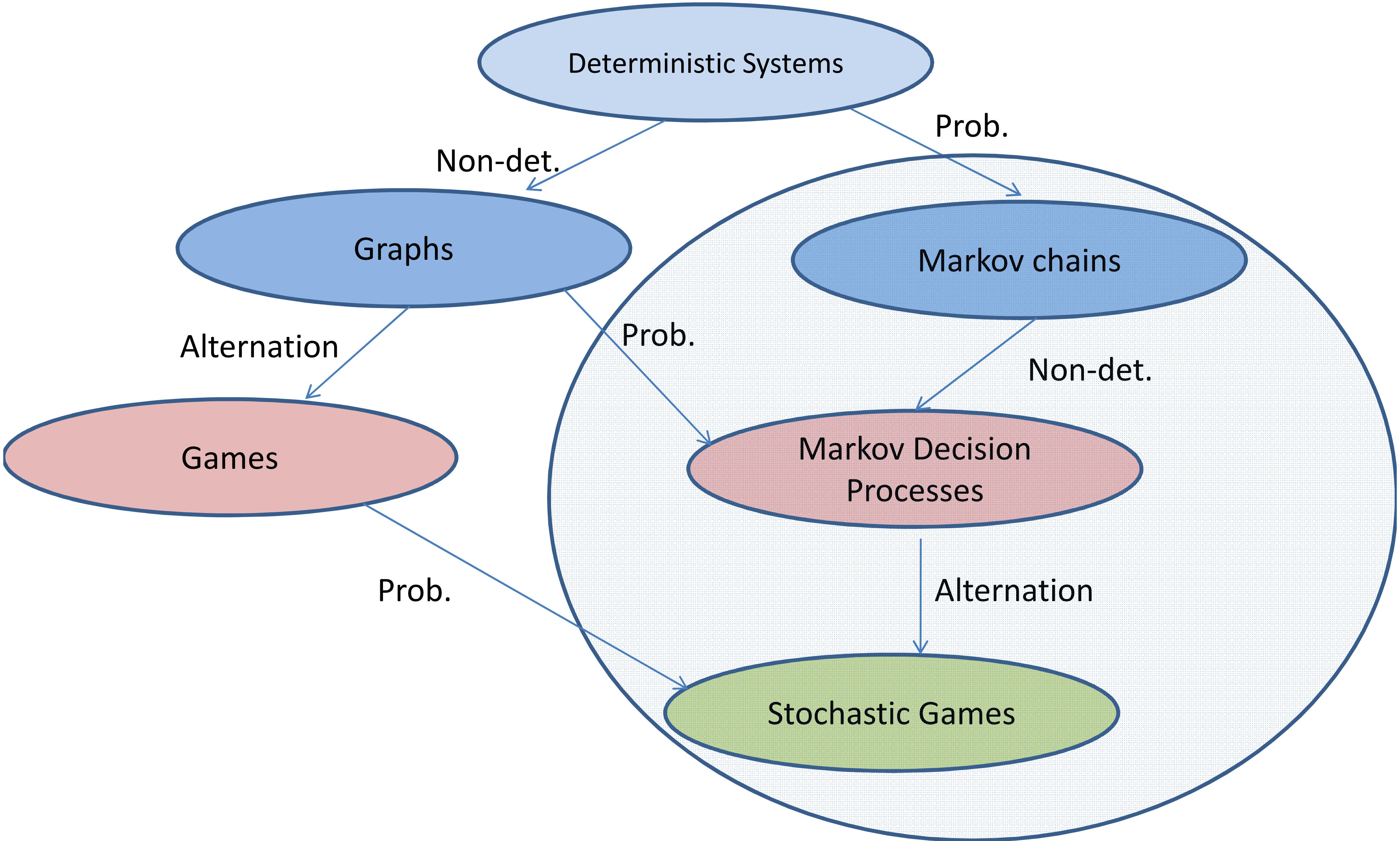
- System to game graph
  - Vertices represent states.
  - Edges represent transitions.
  - Paths represent behavior.
  - Players represent various interacting agents.

- Mathematical framework for system analysis.

Stochastic Games

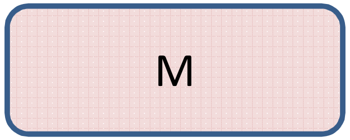


Stochastic Games



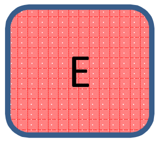
Applications: Verification of Systems

- Verification of systems

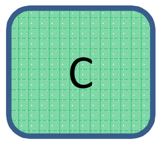


satisfies property

- Environment



- Controller (Synthesis)



Applications: Verification of Systems

- Verification and synthesis of systems
  - System is fixed and the environment fixed: deterministic systems.
  - System is fixed, but not the environment: Demonic non-determinism.
  - Environment fixed but probabilistically (randomized scheduler): Markov chain.
  - Probabilistic environment and controller: Markov decision process.
  - Controller vs. environment: angelic vs. demonic non-determinism (alternation).

Applications: Systems for Specification

- Synthesis of systems from specification
  - Input/Output signals.
  - Automata over I/O that specifies the desired set of behaviors.
  - Can the input player present input such that no matter how the output player plays the generated sequence of I/O signals is accepted by automata ?
  - Deterministic automata: Games.
  - Some input signals generate probabilistic transition: Stochastic games.

Game Models Applications

-synthesis [Church, Ramadge/Wonham, Pnueli/Rosner]

-model checking of open systems

-receptiveness [Dill, Abadi/Lamport]

-semantics of interaction [Abramsky]

-non-emptiness of tree automata [Rabin, Gurevich/ Harrington]

-behavioral type systems and interface automata [deAlfaro/ Henzinger]

-model-based testing [Gurevich/Veanes et al.]

-etc.

• Mathematicians (logic and set theory), Stochastic game theorists, Economists, Computer Scientists, Biologists (evolutionary games).

Properties

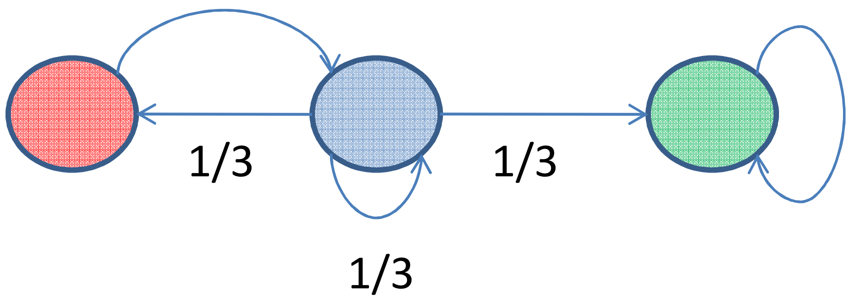
- Properties in verification
  - Reachability to target set.
  - Liveness (Buechi) or repeated reachability.
  - Fairness.
  - Parity objectives: all ω-regular specifications.

MARKOV CHAINS

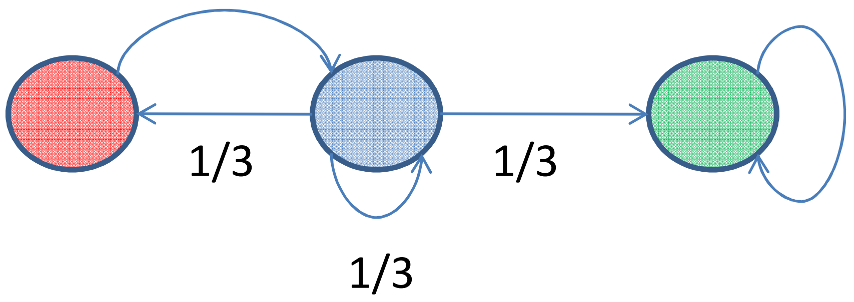
Markov Chains

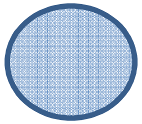
- Markov chain model: $G=((S,E), \delta)$
- Finite set S of states.
- Probabilistic transition function δ
- $E = \{ (s,t) \mid \delta(s)(t) > 0\}$
- The graph (S,E) is useful.

Markov Chain: Example

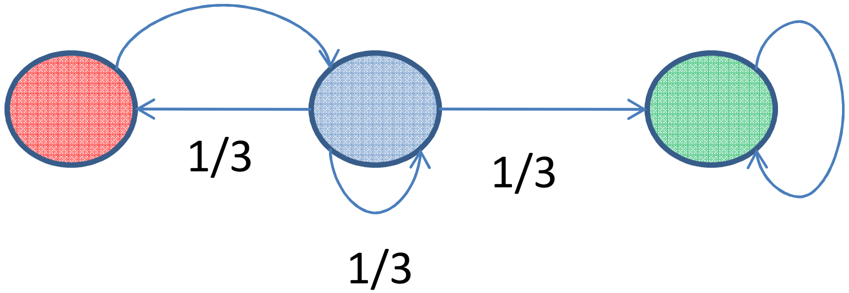


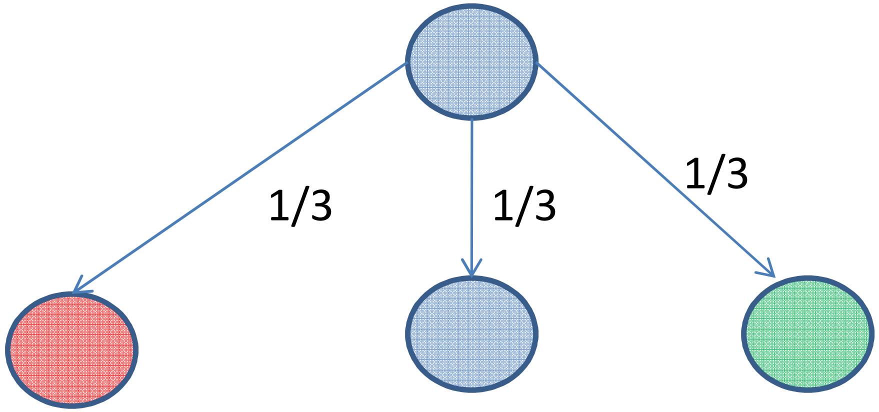
Markov Chain: Example



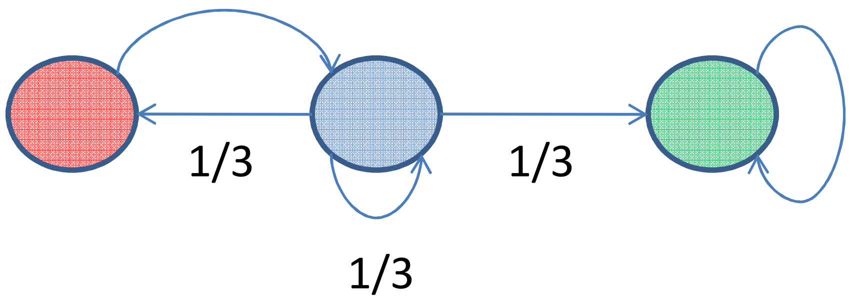


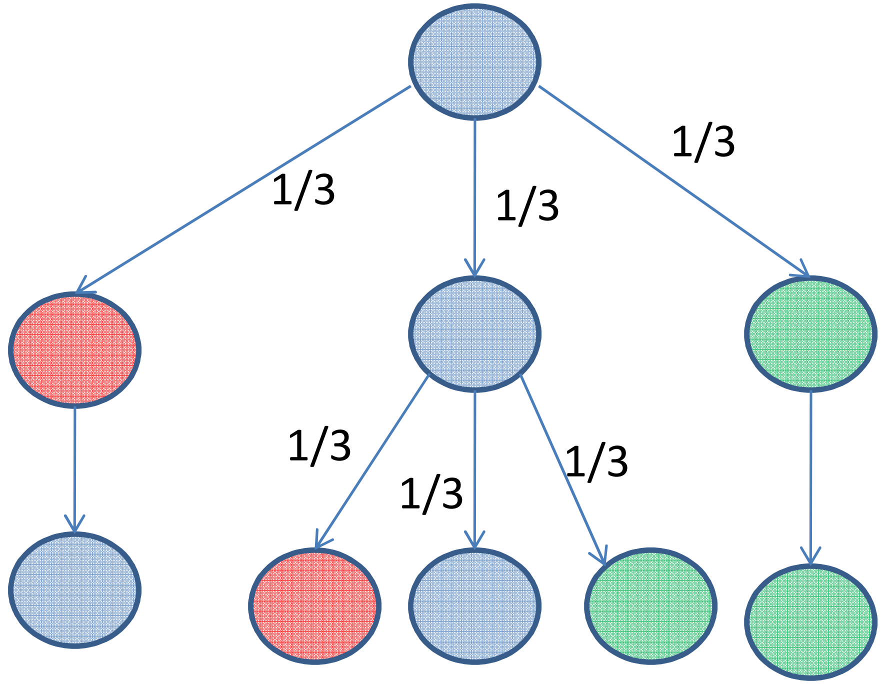
Markov Chain: Example



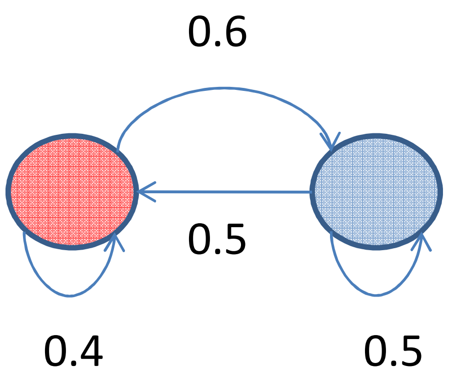


Markov Chain: Example





Markov Chain: Example

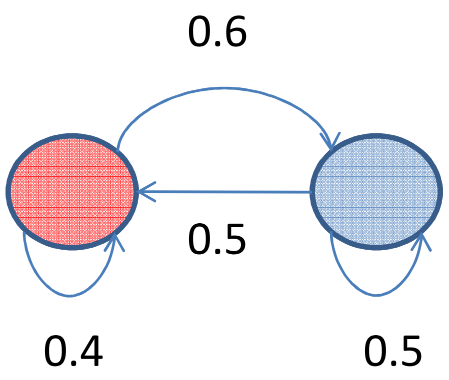


Cola and Pepsi:

Drink Cola today:
Pepsi with prob. 0.6,
Cola with prob. 0.4

Drink Pepsi today:
Pepsi with prob. 0.5
Cola with prob. 0.5

Markov Chain: Example



Drink Cola today:
Pepsi with prob. 0.6,
Cola with prob. 0.4

Drink Pepsi today:
Pepsi with prob. 0.5
Cola with prob. 0.5

Strongly connected Markov chain:
Average frequency.

Linear equations: for every state s we have

$x_s = \sum_t x_t \cdot \delta(t)(s)$

Markov Chain

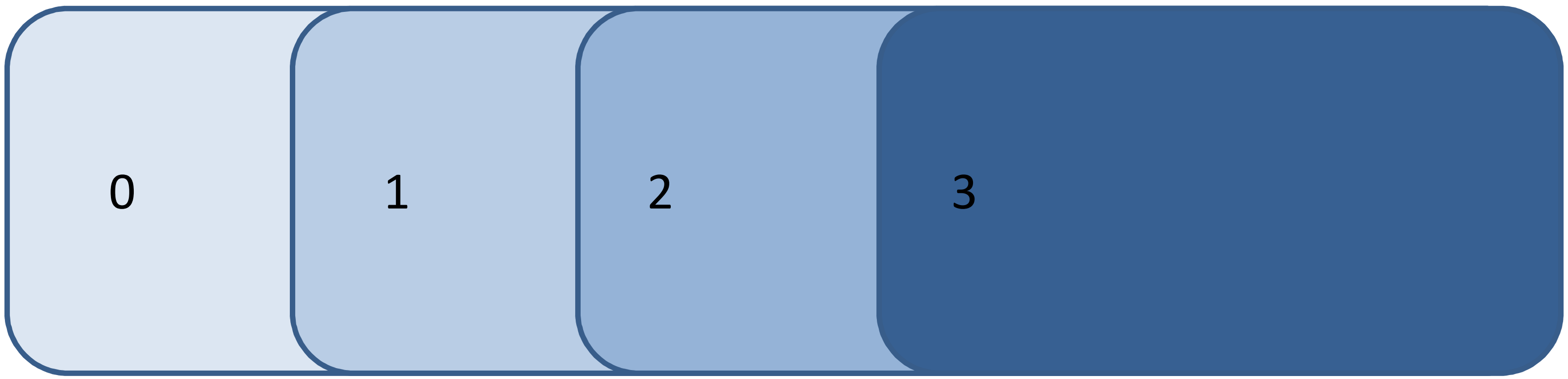
- Properties of interest
  - Target set T: probability to reach the target set.
  - Target set B: probability to visit B infinitely often.

Objectives

- Objectives are subsets of infinite paths, i.e., $\psi \subseteq S^{\omega}$.
- Reachability: set of paths that visit the target T at least once.
- Liveness (Buechi): set of paths that visit the target B infinitely often.
- Parity: given a priority function p: $S \rightarrow \{0,1,\ldots, d\}$, the objective is the set of infinite paths where the minimum priority visited infinitely often is even.

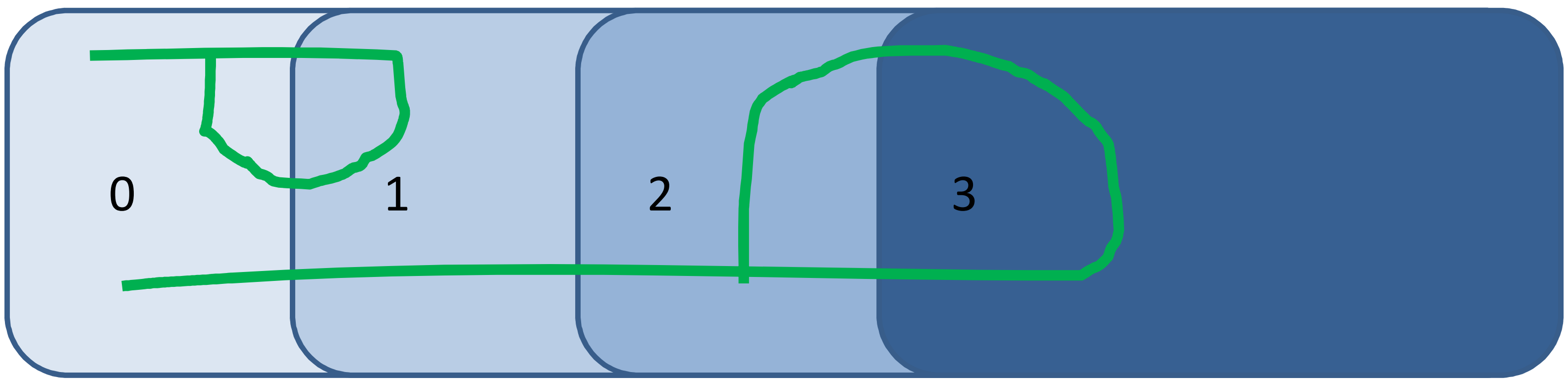
Parity Objectives

- Parity: given a priority function p: S $\rightarrow$ {0,1,…, d}, the objective is the set of infinite paths where the minimum priority visited infinitely often is even.



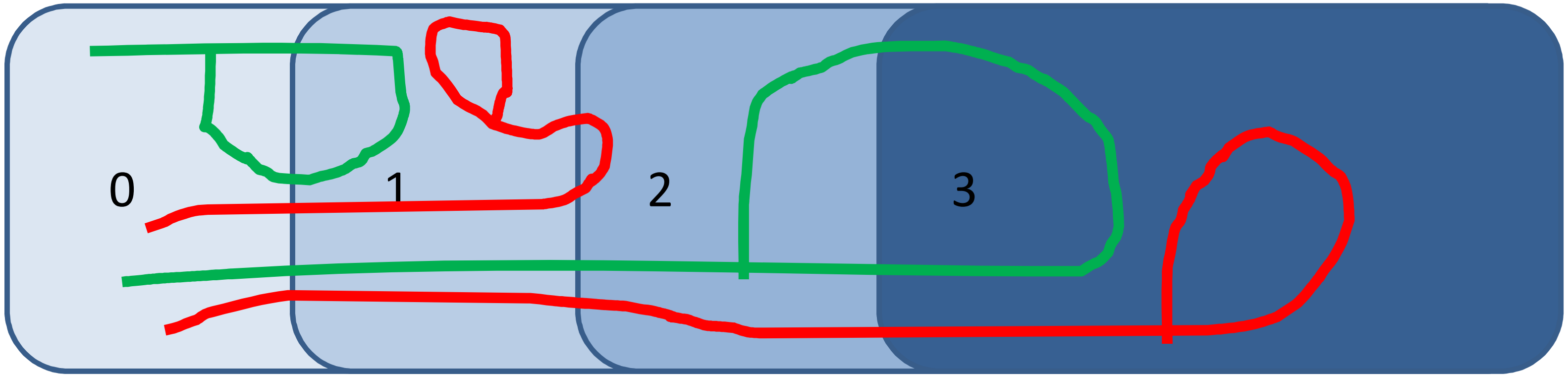
Parity Objectives

- Parity: given a priority function p: S $\rightarrow$ {0,1,…, d}, the objective is the set of infinite paths where the minimum priority visited infinitely often is even.

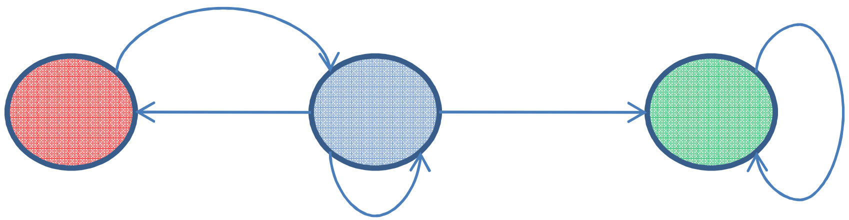


Parity Objectives

- Parity: given a priority function $p: S \rightarrow \{0,1,\ldots, d\}$, the objective is the set of infinite paths where the minimum priority visited infinitely often is even.



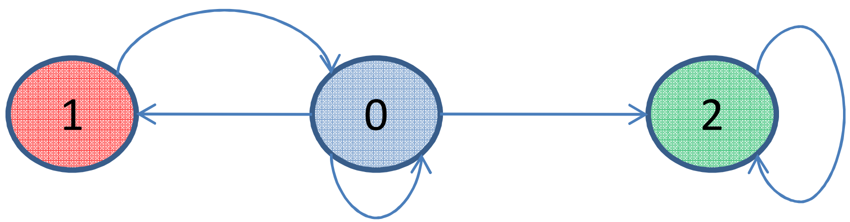
Markov Chain: Example



- Reachability: starting state is blue.
  - Red: probability is less than 1.
  - Blue: probability is 1.
  - Green: probability is 1.

- Liveness: infinitely often visit
  - Red: probability is 0.
  - Blue: probability is 0.
  - Green: probability is 1.

Markov Chain: Example



- Parity
  - Blue infinitely often, or 1 finitely often.
  - In general, if priorities are 0,1, …, 2d, then we require for some $0 \leq i \leq d$, that priority 2i infinitely often, and all priorities less than 2i is finitely often.

Questions

- Qualitative question
  - The set where the property holds with probability 1.
  - Qualitative analysis.

- Quantitative question
  - What is the precise probability that the property holds.
  - Quantitative analysis.

Qualitative Analysis of Markov Chains

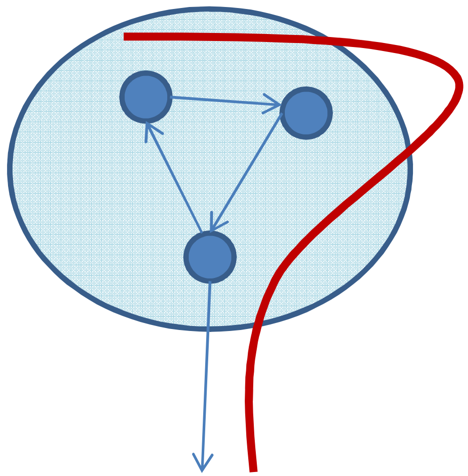
- Consider the graph of Markov chain.
- Closed recurrent set:
  - Bottom strongly connected component.
  - Closed: No probabilistic transition out.
  - Strongly connected.

Qualitative Analysis of Markov Chains

- Theorem: Reach the set of closed recurrent set with probability 1.
- Proof.
  - Consider the DAG of the scc decomposition of the graph.
  - Consider a scc C of the graph that is not bottom.
  - Let α be the minimum positive transition prob.
  - Leave C within n steps with prob at least $\beta = \alpha^n$.
  - Stay in C for at least k*n steps is at most $(1-\beta)^k$.
  - As k goes to infinity this goes to 0.

Qualitative Analysis of Markov Chains

- Theorem: Reach the set of closed recurrent set with probability 1.
- Proof.
  - Path goes out with β.
  - Never gets executed for k times is $(1-\beta)^k$. Now let k goto infinity.



Qualitative Analysis of Markov Chains

- Theorem: Given a closed recurrent set C, for any starting state in C, all states is reached with prob 1, and hence all states visited infinitely often with prob 1.

- Proof. Very similar argument like before.

Qualitative and Quantitative Analysis

- Previous two results are the basis.

- Example: Liveness objective.
  - Compute max scc decomposition.
  - Reach the bottom scc’s with prob 1.
  - A bottom scc with a target is a good bottom scc, otherwise bad bottom scc.
  - Qualitative: if a path to a bad bottom scc, not with prob 1. Otherwise with prob 1.
  - Quantitative: reachability probability to good bottom scc.

Quantitative Reachability Analysis

- Let us denote by C the set of bottom scc's (the quantitative values are 0 or 1). We now define a set of linear equalities. There is a variable x_s for every state s. The equalities are as follows:
  - $x_s = 0$ if s in C and bad bottom scc.
  - $x_s = 1$ if s in C and good bottom scc.
  - $x_s = \sum_{t \in S} x_t * \delta(s)(t)$.
- Brief proof idea: The remaining Markov chain is transient. Matrix algebra $\det(I-\delta) \neq 0$.

Markov Chain Summary

	Reachability	**Liveness**	**Parity**
Qualitative	Linear time	Linear time	Linear time
Quantitative	Linear equalities (Gaussian elimination)	Linear equalities	Linear equalities

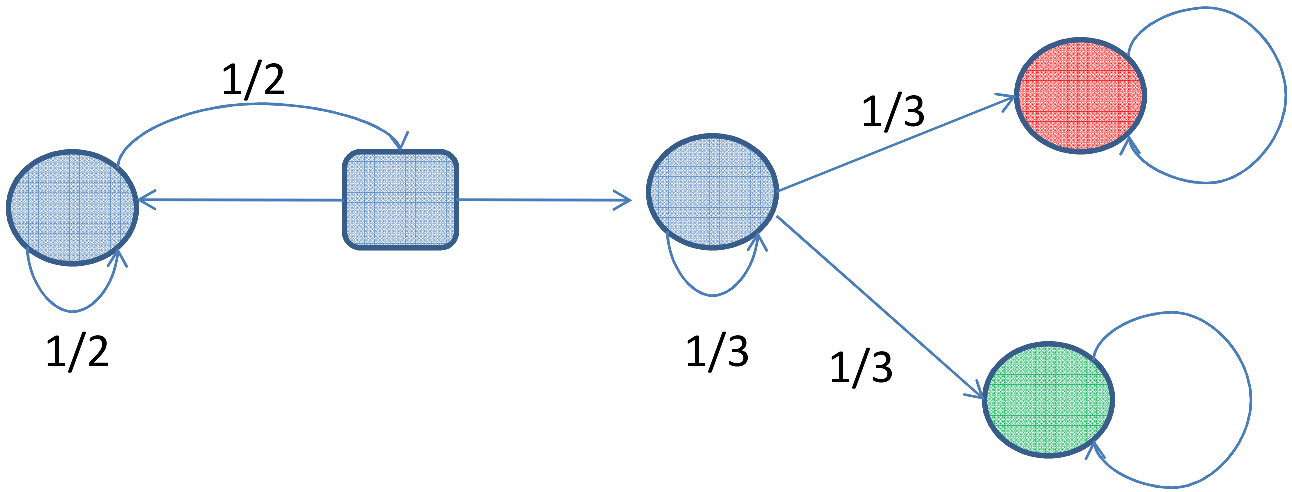
MARKOV DECISION PROCESSES

Markov Decision Processes

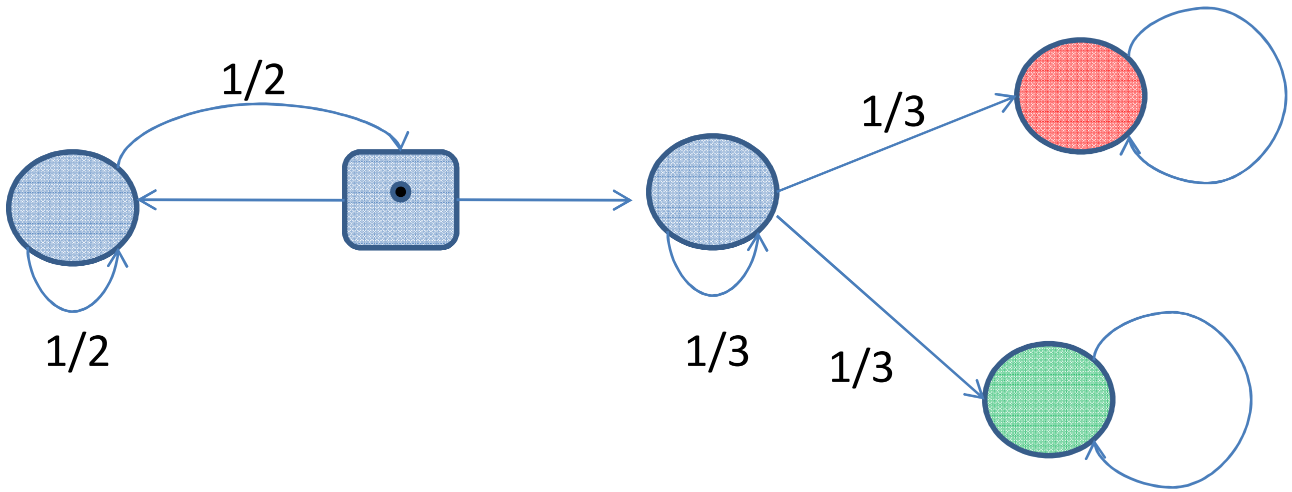
- Markov decision processes (MDPs)
  - Non-determinism.
  - Probability.
  - Generalizes non-deterministic systems and Markov chains.

- An MDP $G = ((S,E), (S_1, S_P), \delta)$
  - $\delta : S_P \to D(S)$.
  - For $s \in S_P$, the edge $(s,t) \in E$ iff $\delta(s)(t)>0$.
  - E(s) out-going edges from s, and assume E(s) non-empty for all s.

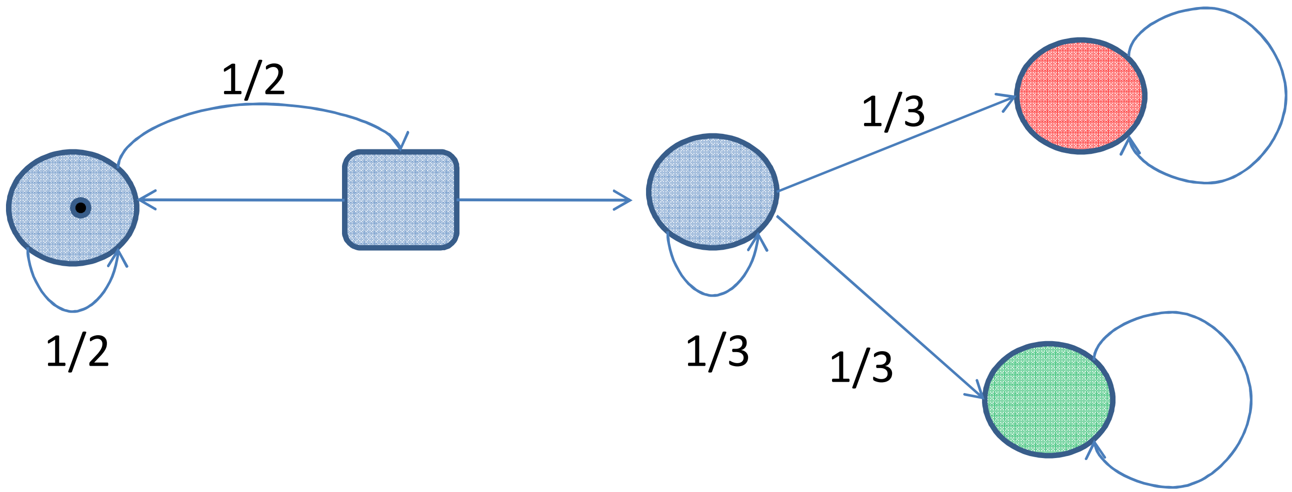
MDP: Example



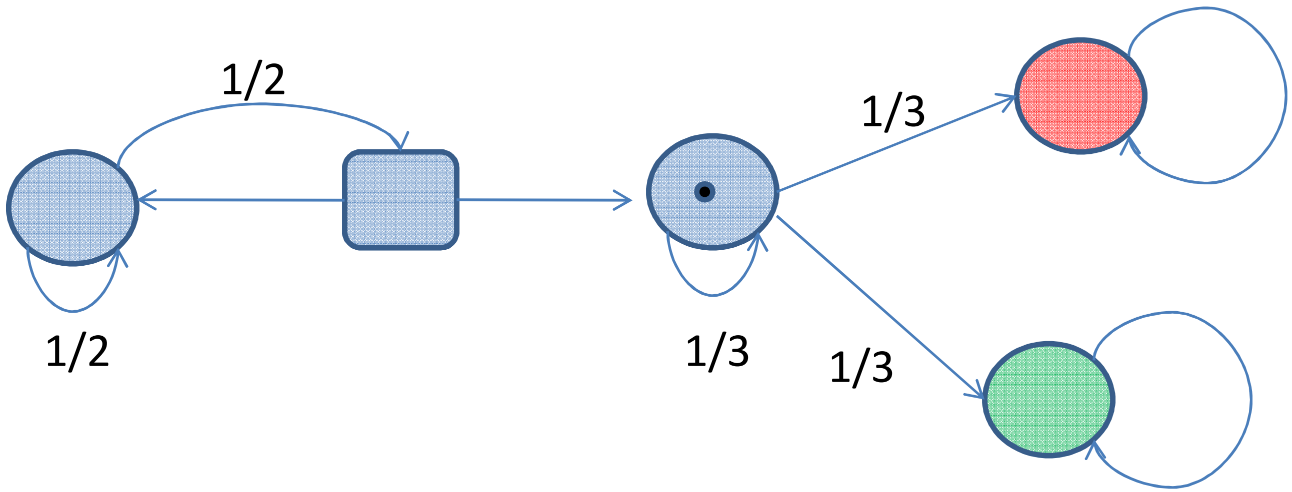
MDP: Example



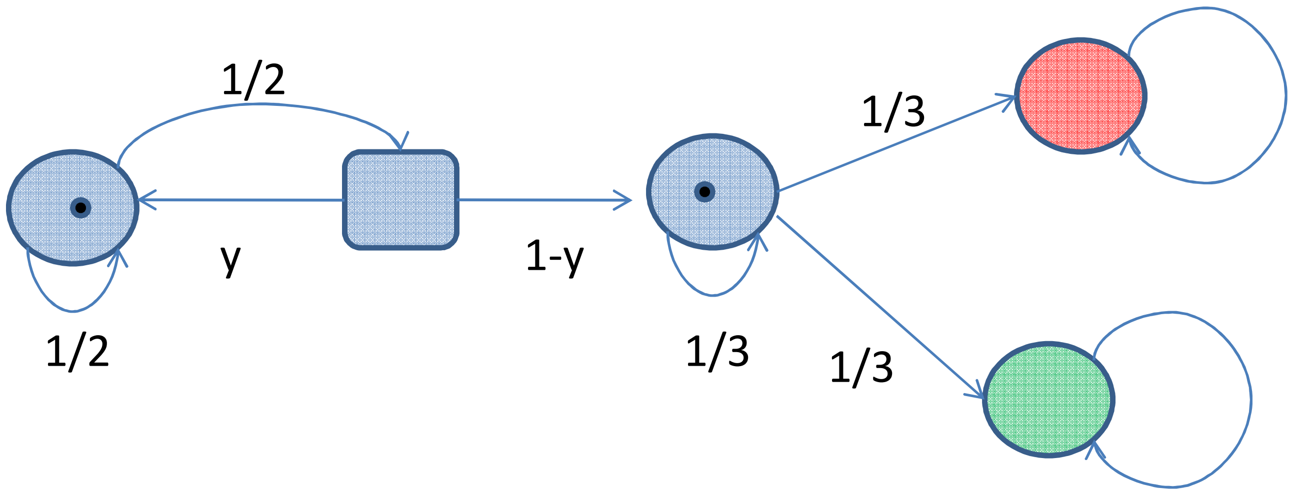
MDP: Example



MDP: Example



MDP: Example



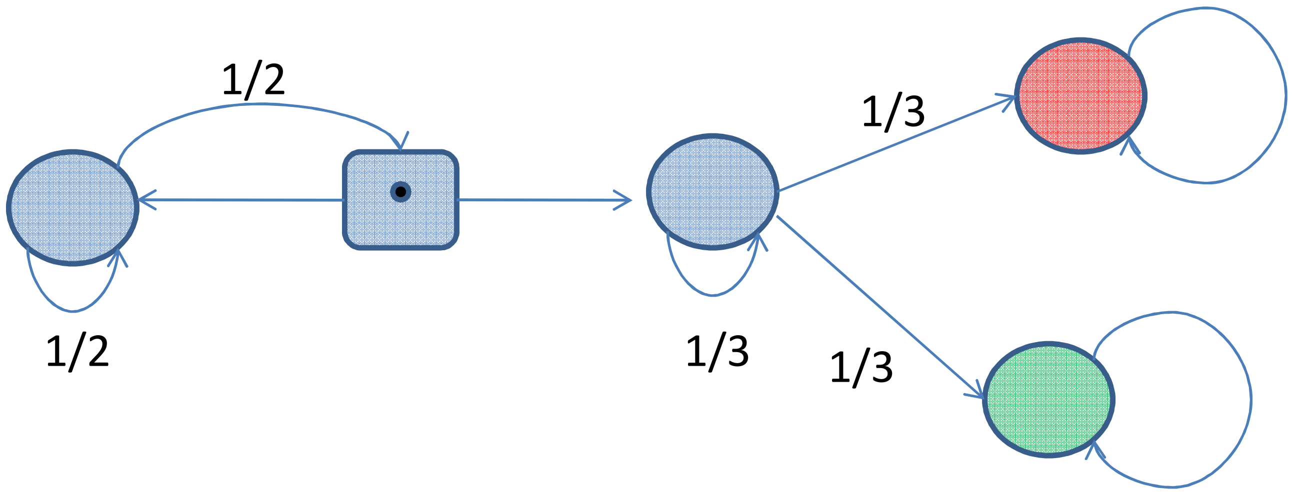
MDP

- Model
- Objectives
- How is non-determinism resolved: notion of strategies. At each stage can be resolved differently and also probabilistically.

Strategies

- Strategies are recipe how to move tokens or how to extend plays. Formally, given a history of play (or finite sequence of states), it chooses a probability distribution over out-going edges.
  - σ: $S^* S_1 \to D(S)$.

MDP: Strategy Example

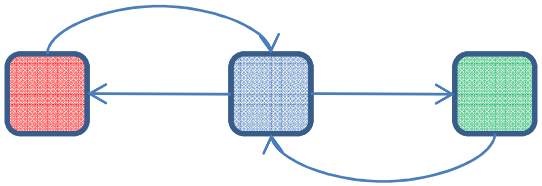


Token for k-th time: choose left with prob 1/k and right (1-1/k).

Strategies

- Strategies are recipe how to move tokens or how to extend plays. Formally, given a history of play (or finite sequence of states), it chooses a probability distribution over out-going edges.
  - σ: $S^* S_1 \rightarrow D(S)$.
- History dependent and randomized.
- History independent: depends only current state (memoryless or positional).
  - σ: $S_1 \rightarrow D(S)$
- Deterministic: no randomization (pure strategies).
  - σ: $S^* S_1 \rightarrow S$
- Deterministic and memoryless: no memory and no randomization (pure and memoryless and is the simplest class).
  - σ: $S_1 \rightarrow S$

Example: Cheating Lovers



Visit green and red infinitely often.

Pure memoryless not good enough.

Strategy with memory: alternates.

Randomized memoryless: choose with uniform probability.

Certainty vs. probability 1.

Values in MDPs

- Value at a state for an objective ψ
  - $\mathrm{Val}(\psi)(s) = \sup_\sigma \Pr_s^\sigma(\psi)$.
- Qualitative analysis
  - Compute the set of almost-sure (prob 1) winning states (i.e., set of states with value 1).
- Quantitative analysis
  - Compute the value for all states.

Qualitative and Quantitative Analysis

- Qualitative analysis
  - Liveness (Buechi) and reachability as a special case.
- Reduction of quantitative analysis to quantitative reachability.
- Quantitative reachability.

Qualitative Analysis for Liveness

- An MDP G, with a target set B.
- Set of states such that there is a strategy to ensure that B is visited infinitely often with probability 1.
- We will show pure memoryless is enough.
- The generalization to parity (left as an exercise).

Attractor

- Random Attractor for a set U of states.
- $U_0 = U$.
- $U_{i+1} = U_i \cup \{s \in S_1 \mid E(s) \subseteq U_i\}$
  $\cup \{s \in S_P \mid E(s) \cap U_i \neq \emptyset\}$.
- From U_{i+1} no matter what is the choice, U_i is reached with positive probability. By induction U is reached with positive probability.

Attractor

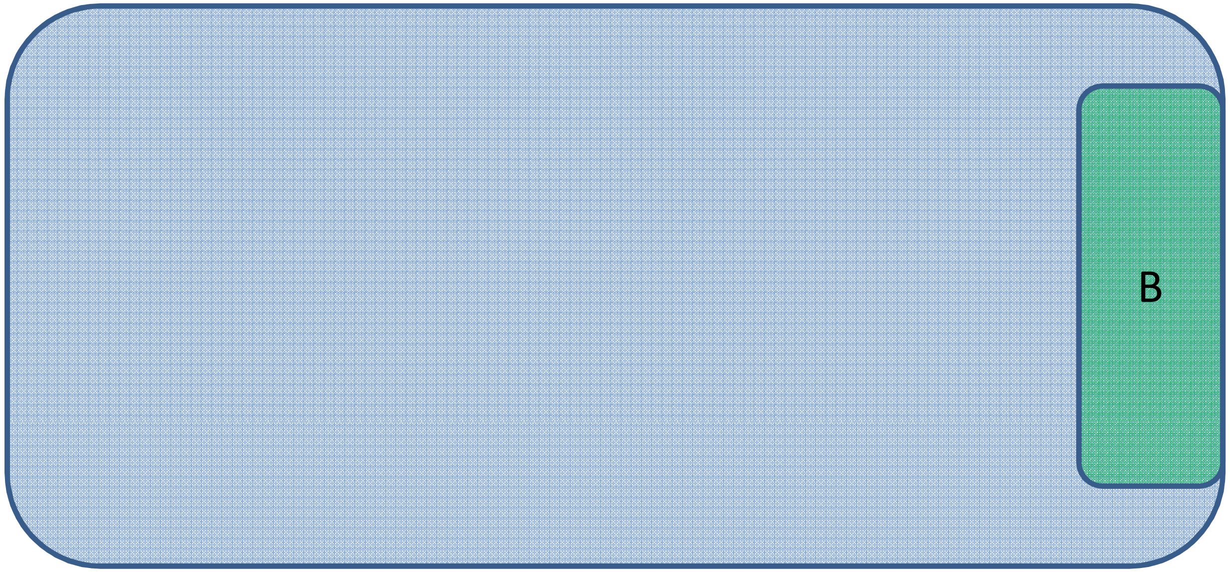
- $Attr_P(U) = \cup_{i \geq 0} U_i$.

- Attractor lemma: From $Attr_P(U)$ no matter the strategy of the player (history dependent, randomized) the set U is reached with positive probability.

- Can be computed in O(m) time (m number of edges).

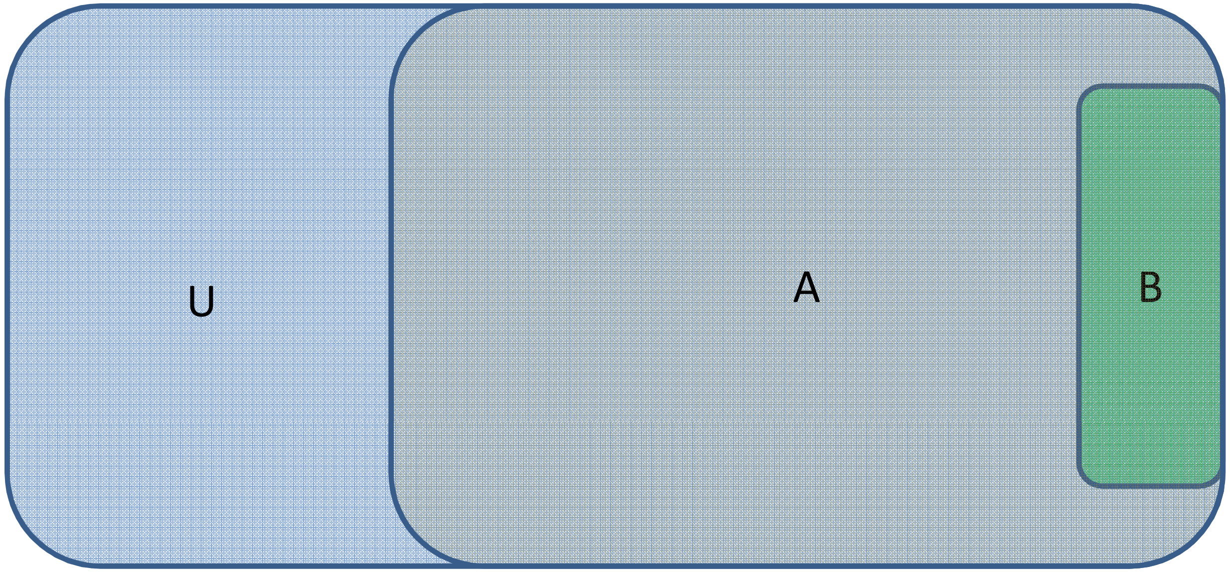
- Thus if U is not in the almost-sure winning set, then $Attr_P(U)$ is also not in the almost-sure winning set.

Iterative Algorithm



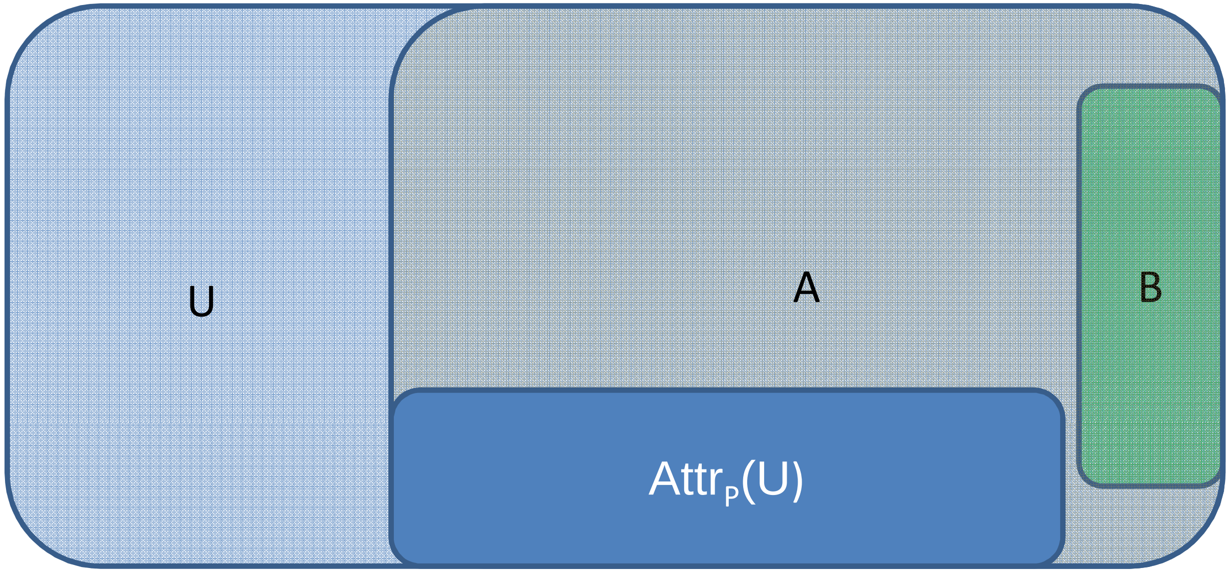
- Compute simple reachability to B (exist a path in the graph of the MDP (S,E). Let us call this set A.

Iterative Algorithm



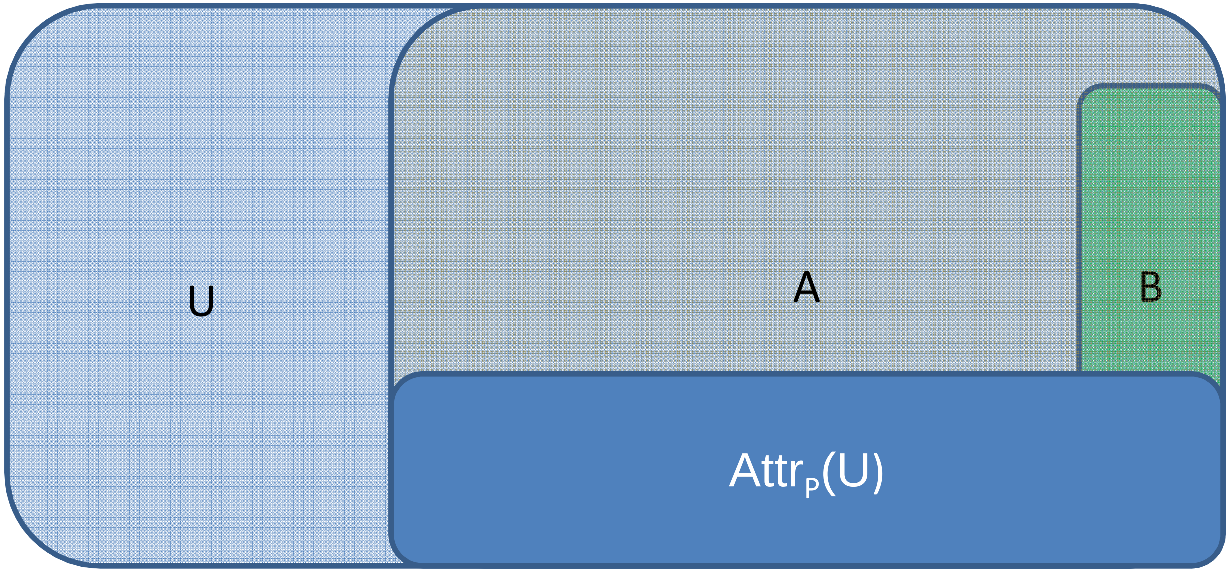
- Let U= S \ A. Then there is not even a path from U to B. Clearly, U is not in the almost-sure set.
- By attractor lemma can take $Attr_P(U)$ out and iterate.

Iterative Algorithm



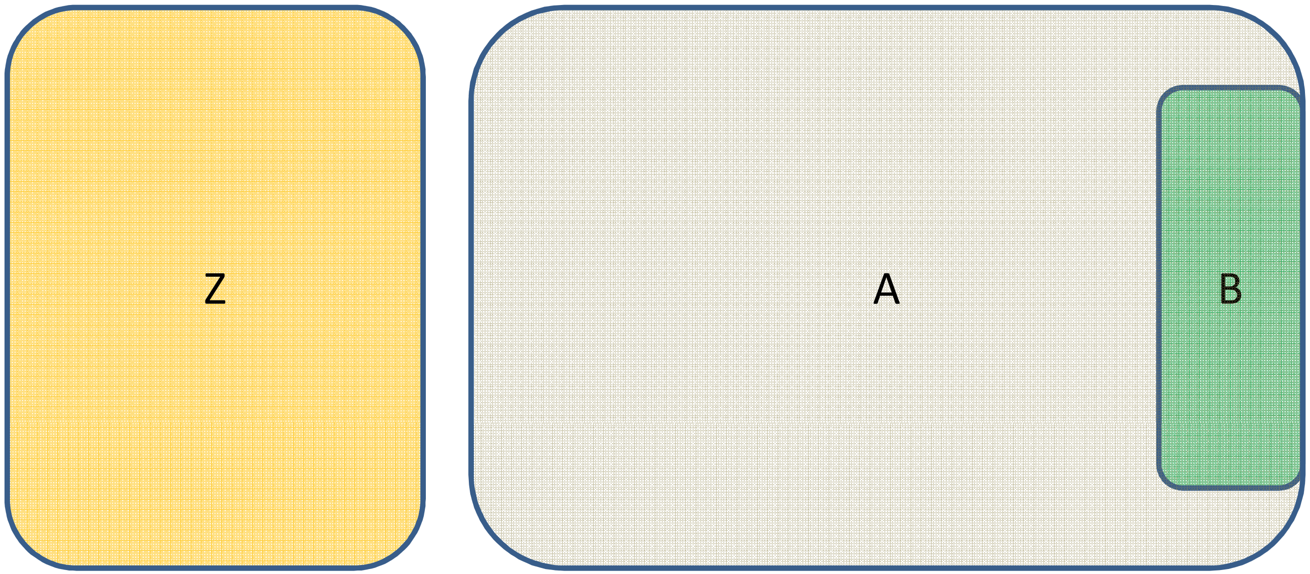
- $Attr_P(U)$ may or may not intersect with B.

Iterative Algorithm



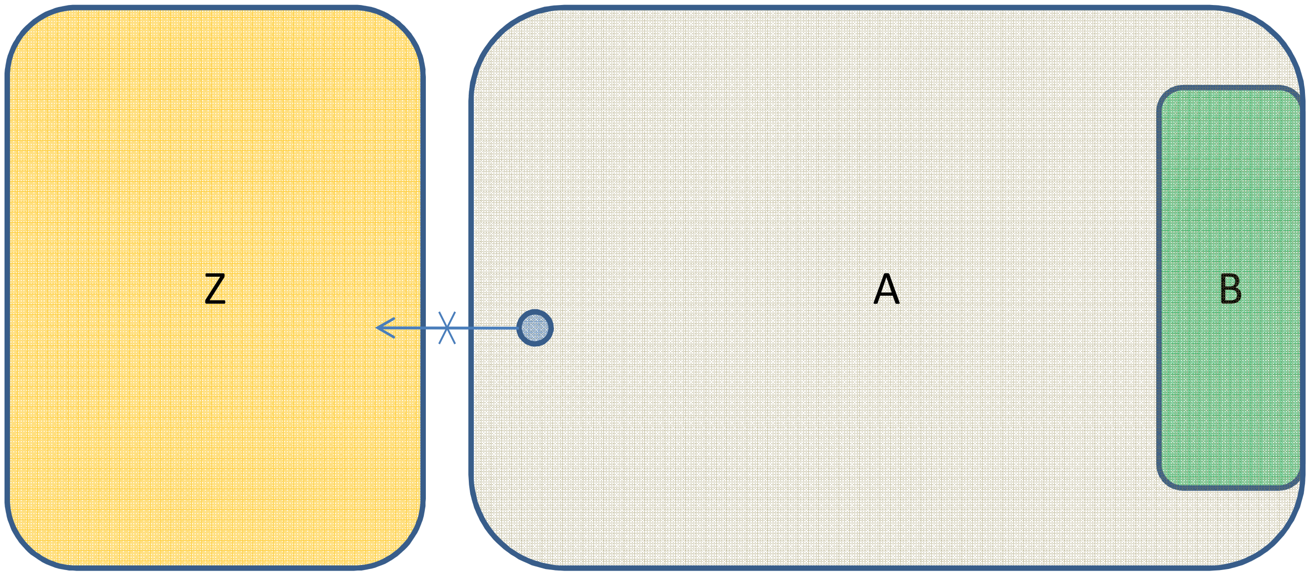
- Iterate on the remaining sub-graph.
- Every iteration what is removed is not part of almost-sure winning set.
- What happens when the iteration stops.

Iterative Algorithm



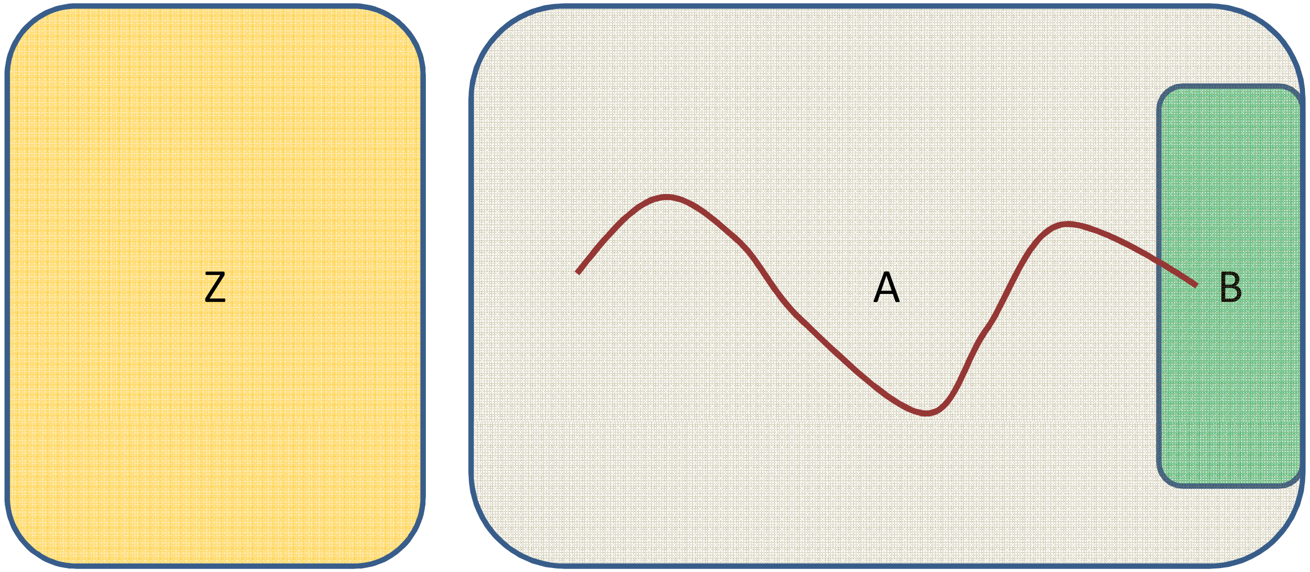
- The iteration stops. Let Z be the set of states removed overall iteration.
- Two key properties.

Iterative Algorithm



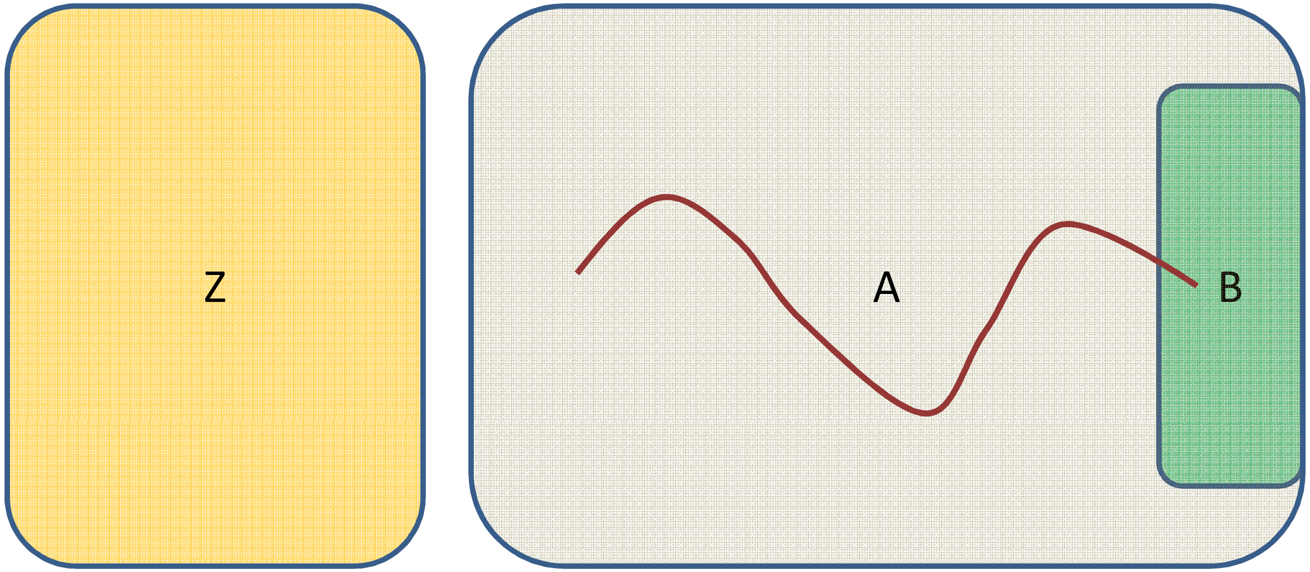
- The iteration stops. Let Z be the set of states removed overall iteration.
- Two key properties:
  - No probabilistic edge from outside to Z.

Iterative Algorithm



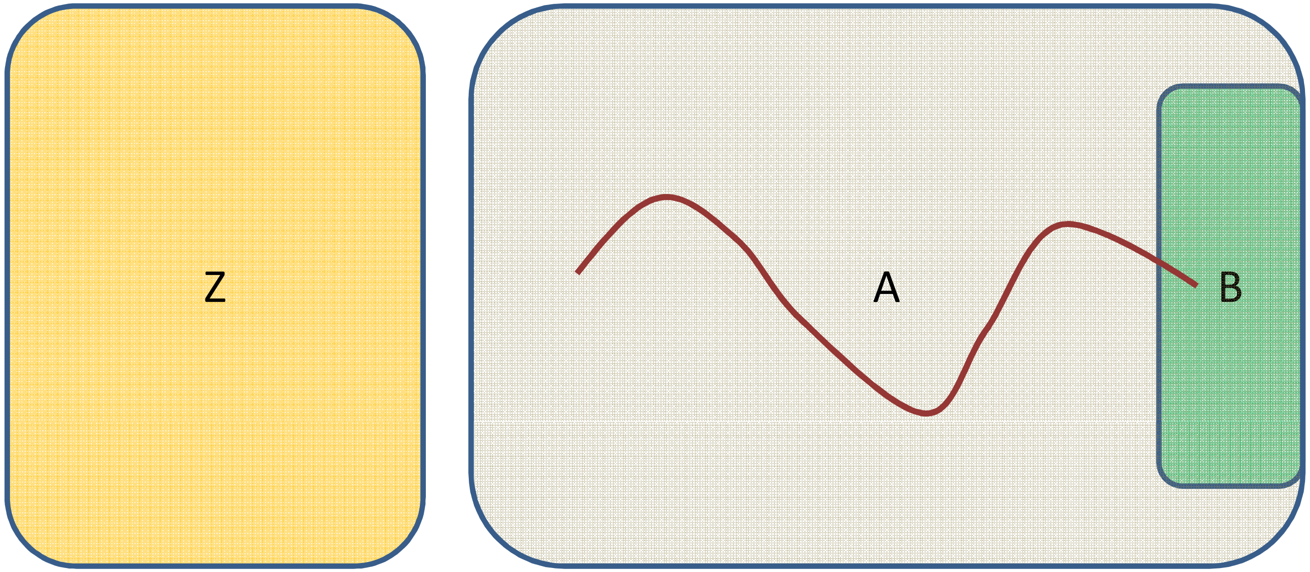
- The iteration stops. Let Z be the set of states removed overall iteration.
- Two key properties:
  - No probabilistic edge from outside to Z.
  - From everywhere in A (the remaining graph) path to B.

Iterative Algorithm



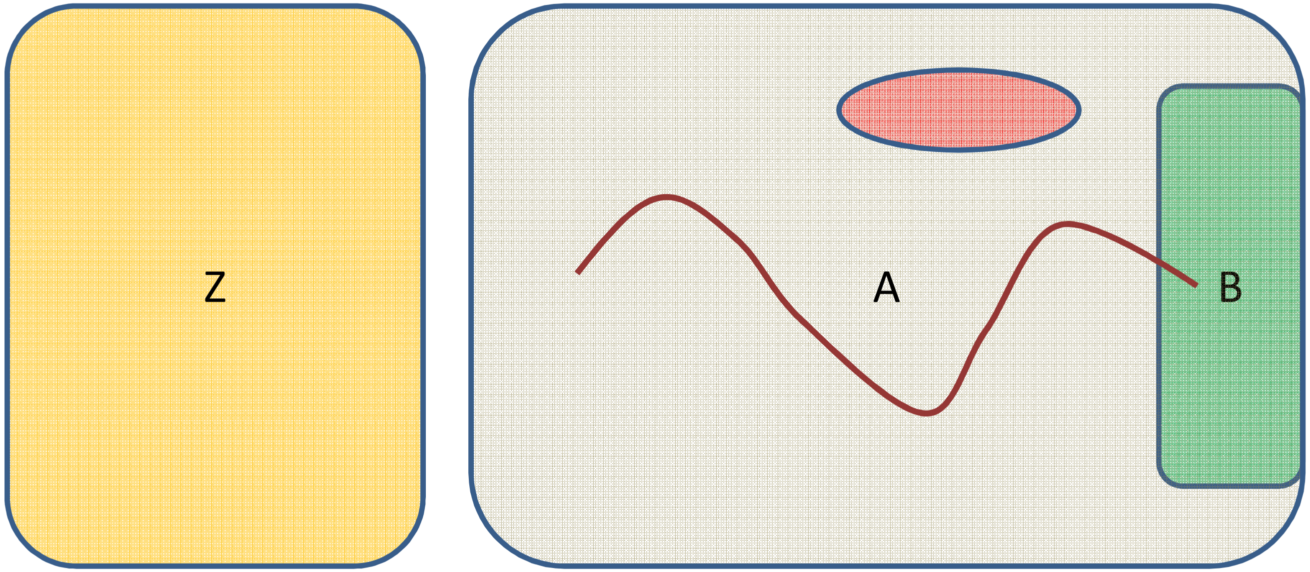
- Two key properties:
  - No probabilistic edge from outside to Z.
  - From everywhere in A (the remaining graph) path to B.
- Fix a memoryless strategy as follows:
  - In $A \setminus B$: shorten distance to B. (Consider the BFS and choose edge).
  - In B: stay in A.

Iterative Algorithm



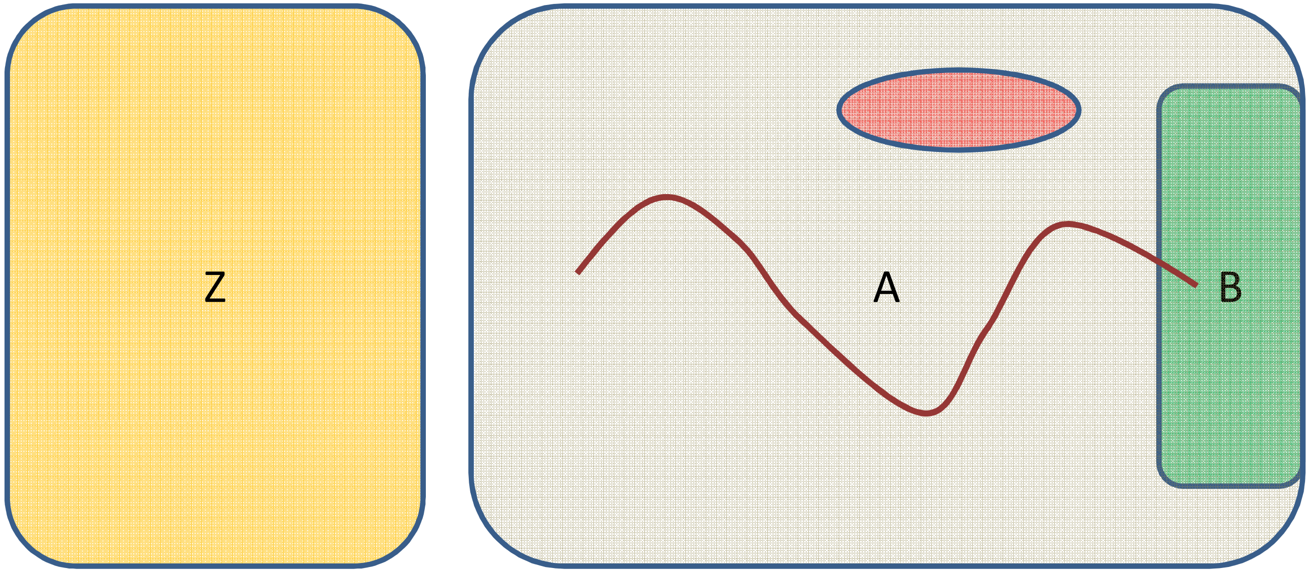
- Fix a memoryless strategy as follows:
  - In A \ B: shorten distance to B. (Consider the BFS and choose edge).
  - In B: stay in A.
- Argue all bottom scc’s intersect with B. By Markov chain theorem done.

Iterative Algorithm



- Argue all bottom scc's intersect with B. By Markov chain theorem done.
- Towards contradiction some bottom scc that does not intersect.
  - Consider the minimum BFS distance to B.

Iterative Algorithm



- Argue all bottom scc's intersect with B. By Markov chain theorem done.
- Towards contradiction some bottom scc that does not intersect.
  - Consider the minimum BFS distance to B.
    - Case 1: if a state in S_P, all edges must be there and so must be the one with shorter distance.
    - Case 2: if a state in S_1, then the successor chosen has shorter distance.
    - In both cases we have a contradiction.

Iterative Algorithm

- Time complexity is O(n m).
- Pure memoryless almost-sure winning strategy.
- Exercise: extend it to parity with time complexity O(n m d).
- We are now done with qualitative analysis. We will now argue how to reduce quantitative analysis to quantitative reachability.

End of Part 1:

1. Markov chains:
   Qualitative and quantitative Analysis

2. MDPs:
   Qualitative analysis

Next Part:

1. MDPs:
   Quantitative Analysis

2. Stochastic games:
   Qualitative and quantitative Analysis