Stochastic Games

Krishnendu Chatterjee

I|S|T AUSTRIA
Institute of Science and Technology

MoVeP, Aachen
June 28, 2010

Stochastic Games

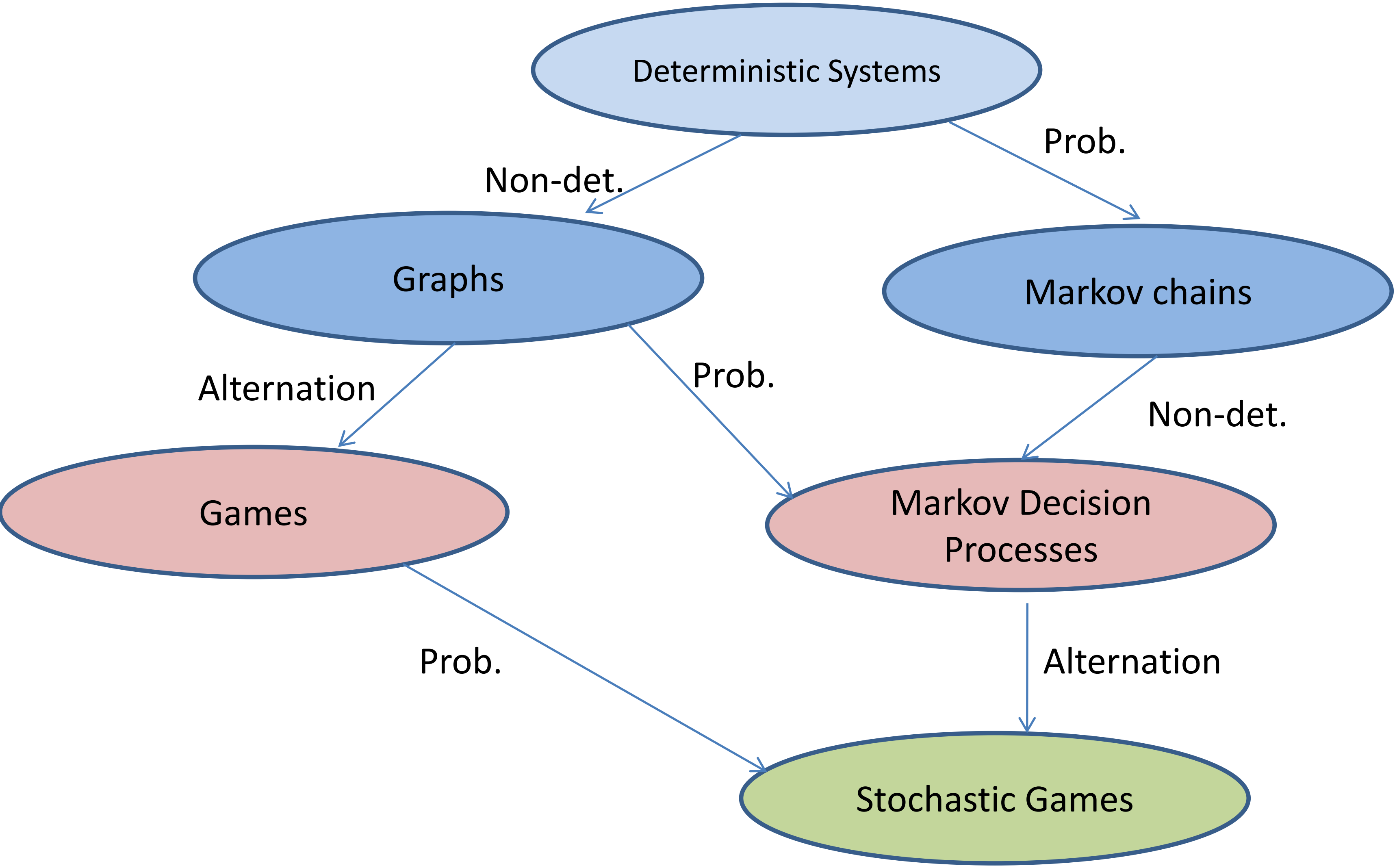
- Two-player perfect-information games on graphs with randomness in transitions.

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

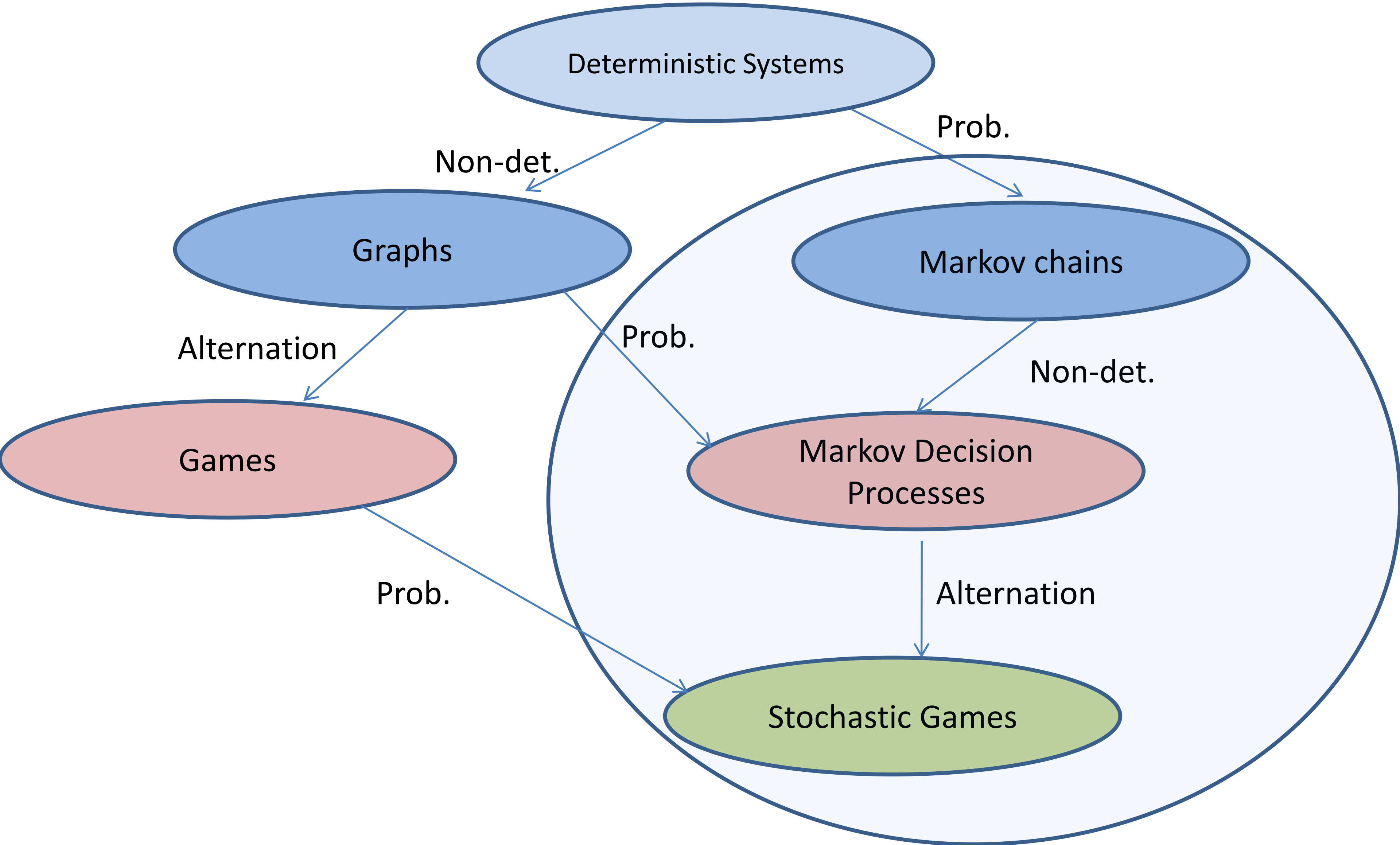
System Analysis

- 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

M

satisfies property

- Environment

E

- Controller (Synthesis)

C

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

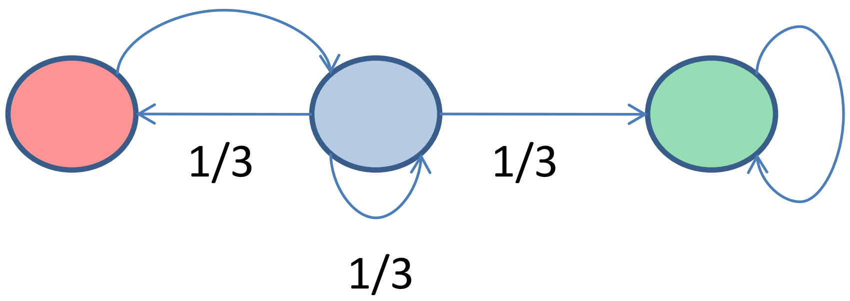
- Properties in verification
  - Reachability to target set.
  - Liveness (Buechi) or repeated reachability.
  - Fairness.
  - Parity objectives: all ω-regular specifications.

MARKOV CHAINS

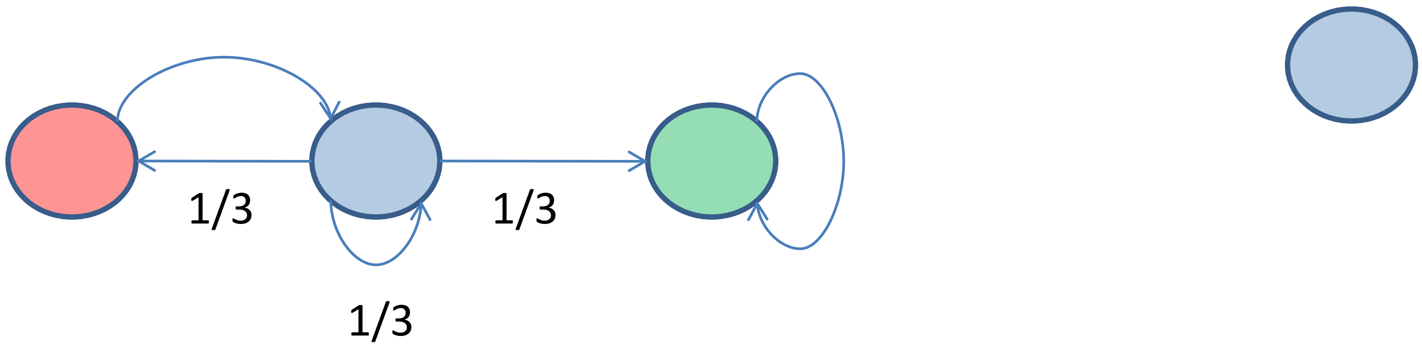
Markov Chains

- 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.

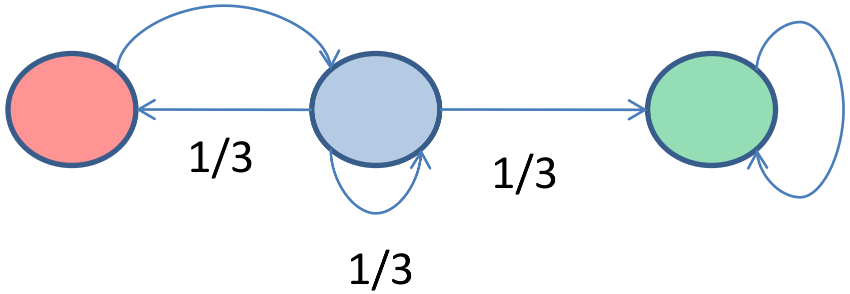
Markov Chain: Example

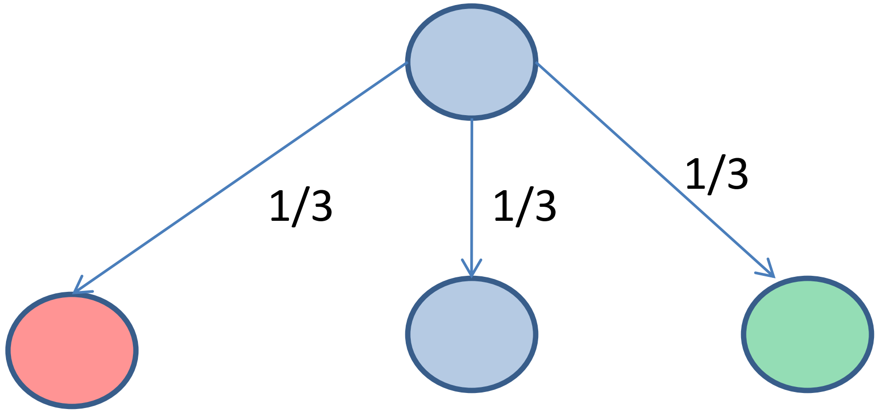


Markov Chain: Example

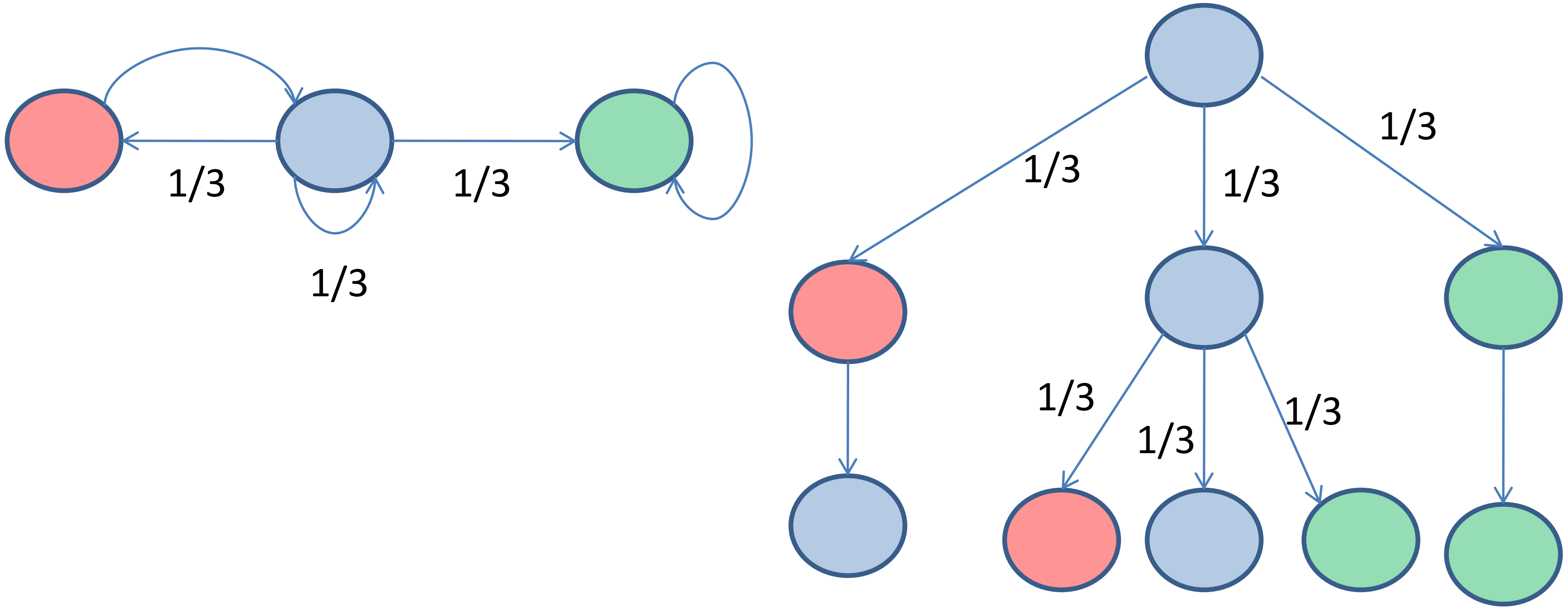


Markov Chain: Example





Markov Chain: Example



Markov Chain

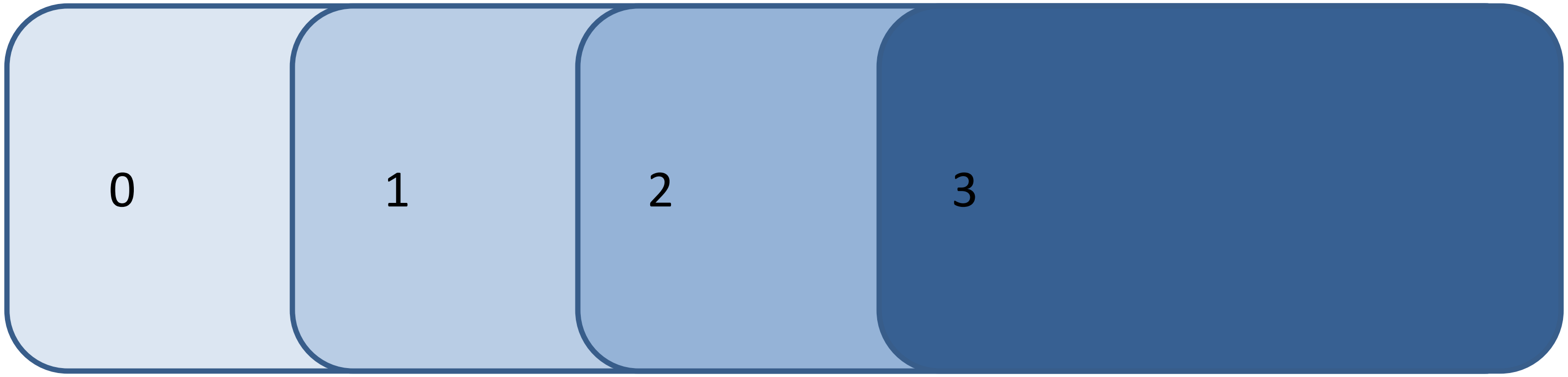
- Properties of interest
  - Target set T: probability to reach the target set.
  - Target set B: probability to visit B infinitely often.

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,…, d}, the objective is the set of infinite paths where the minimum priority visited infinitely often is even.

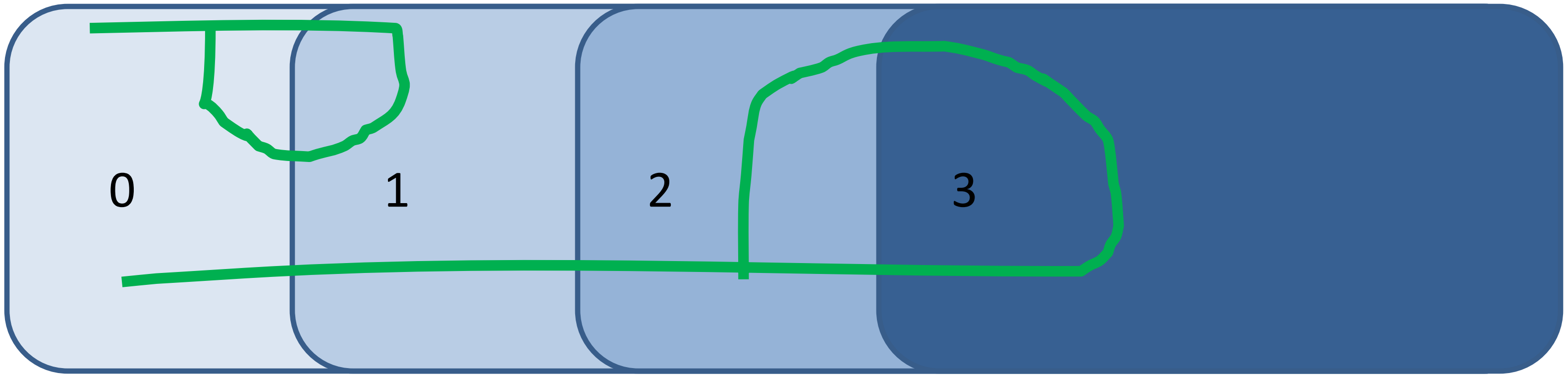
Parity Objectives

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



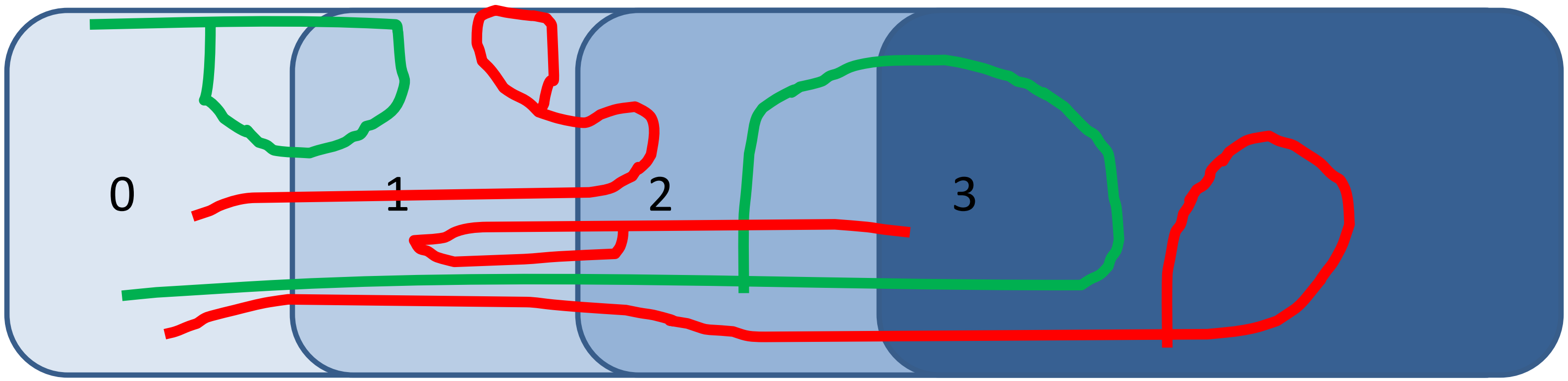
Parity Objectives

- 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.

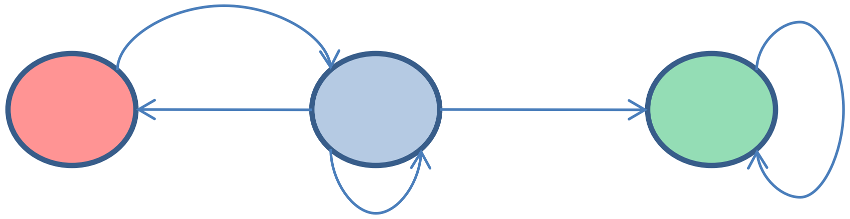


Parity Objectives

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

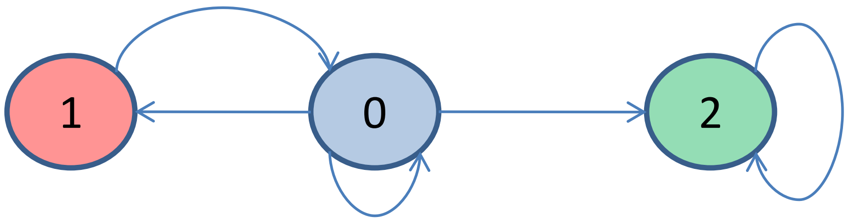


Markov Chain: Example



- 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

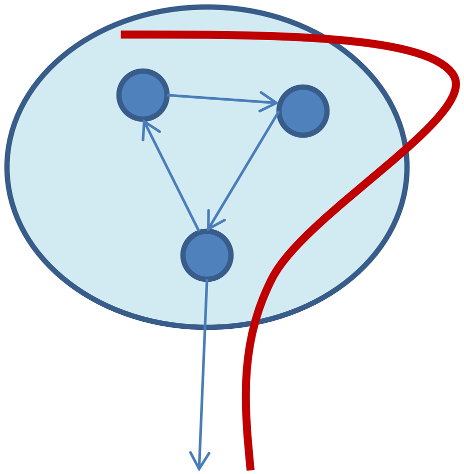
- Consider the graph of Markov chain.
- Closed recurrent set:
  - Bottom strongly connected component.
  - Closed: No probabilistic transition out.
  - Strongly connected.

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

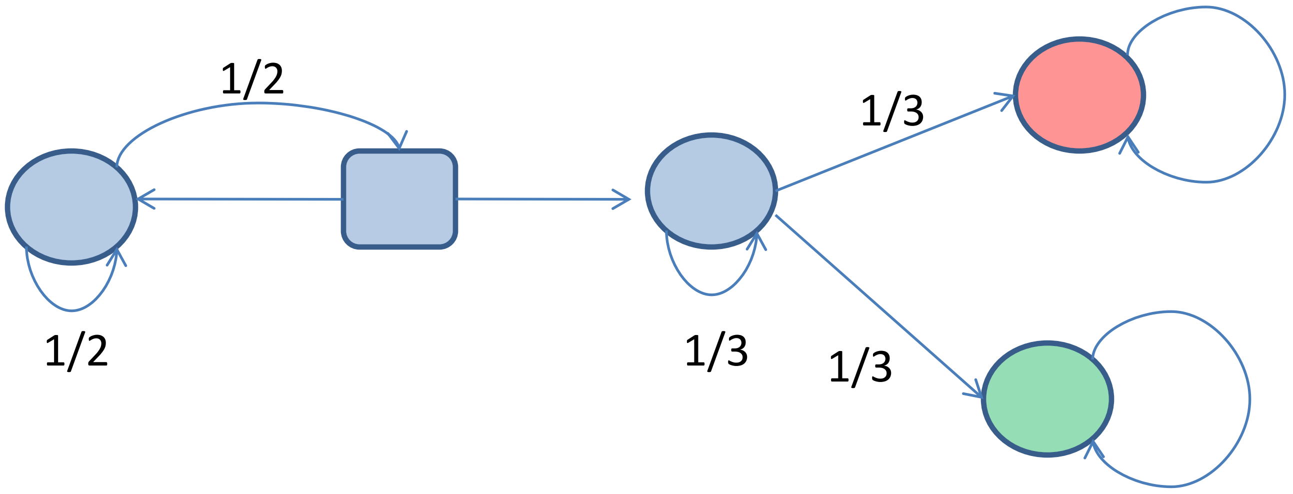
MARKOV DECISION PROCESSES

Markov Decision Processes

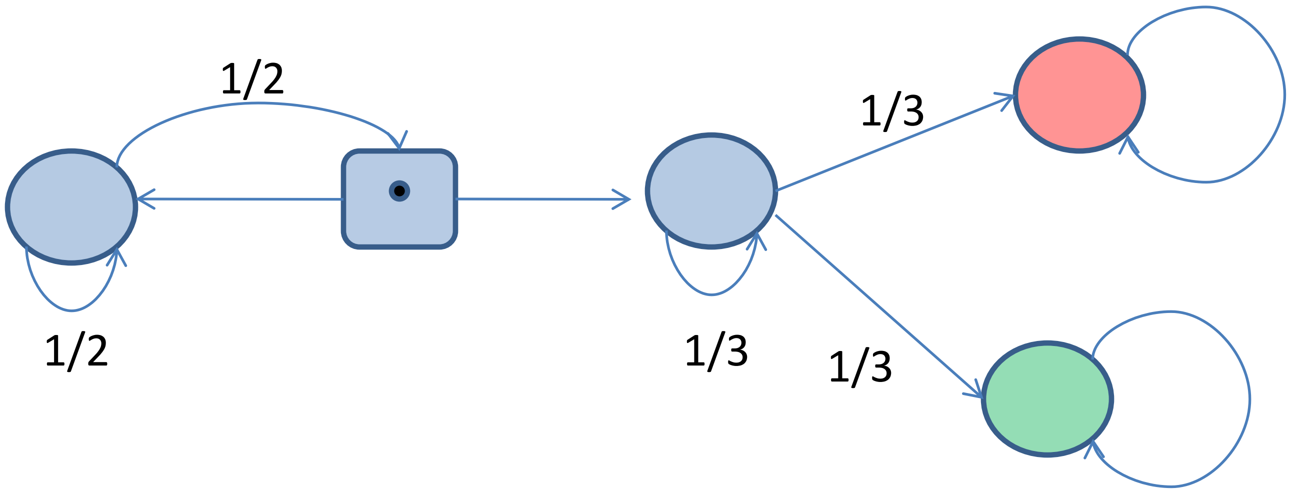
- Markov decision processes (MDPs)
  - Non-determinism.
  - Probability.
  - Generalizes non-deterministic systems and Markov chains.

- 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.

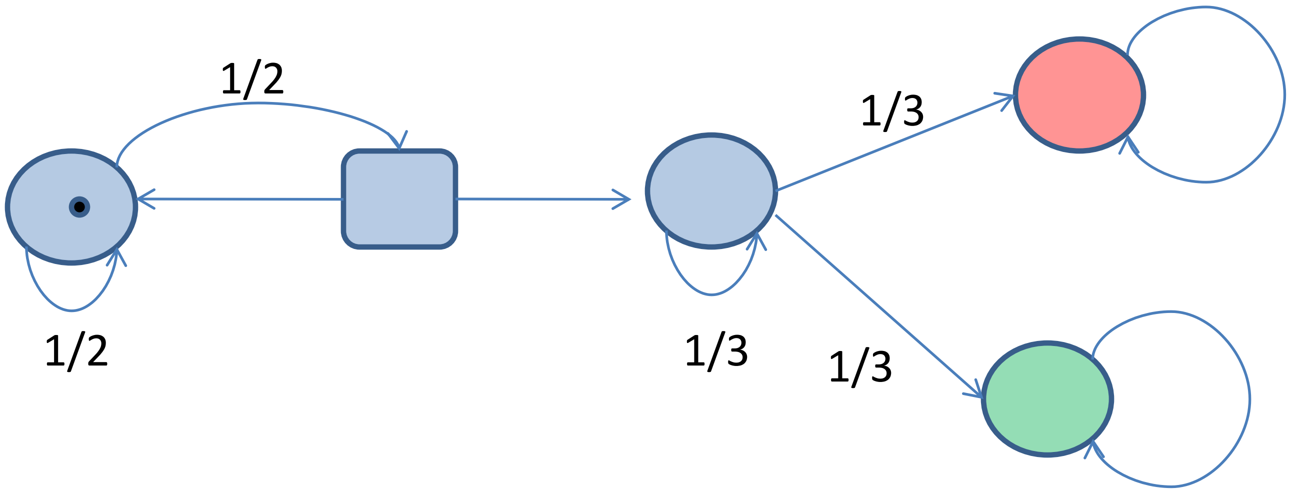
MDP: Example



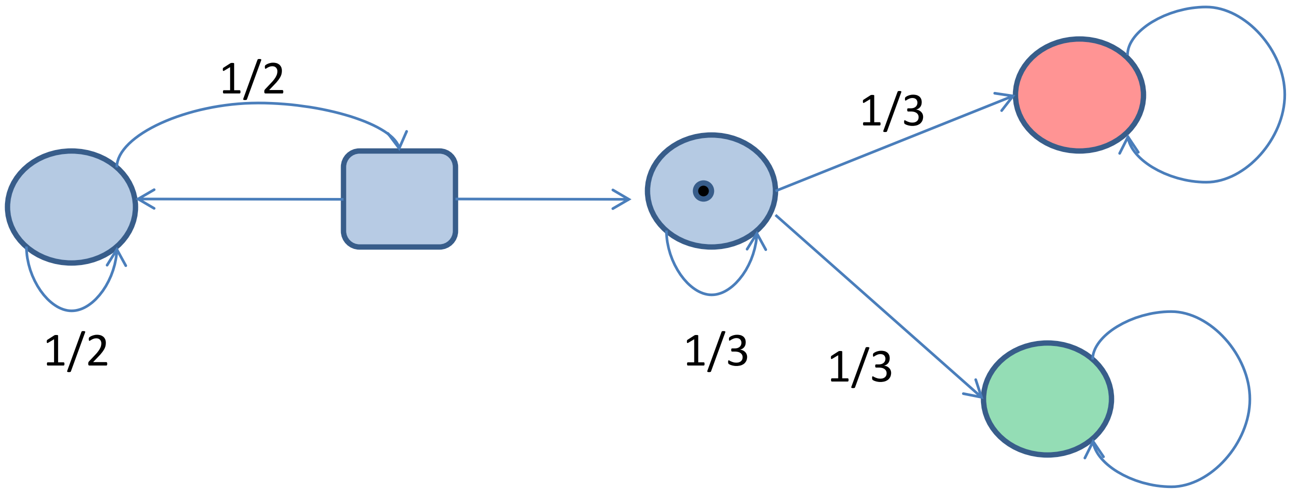
MDP: Example



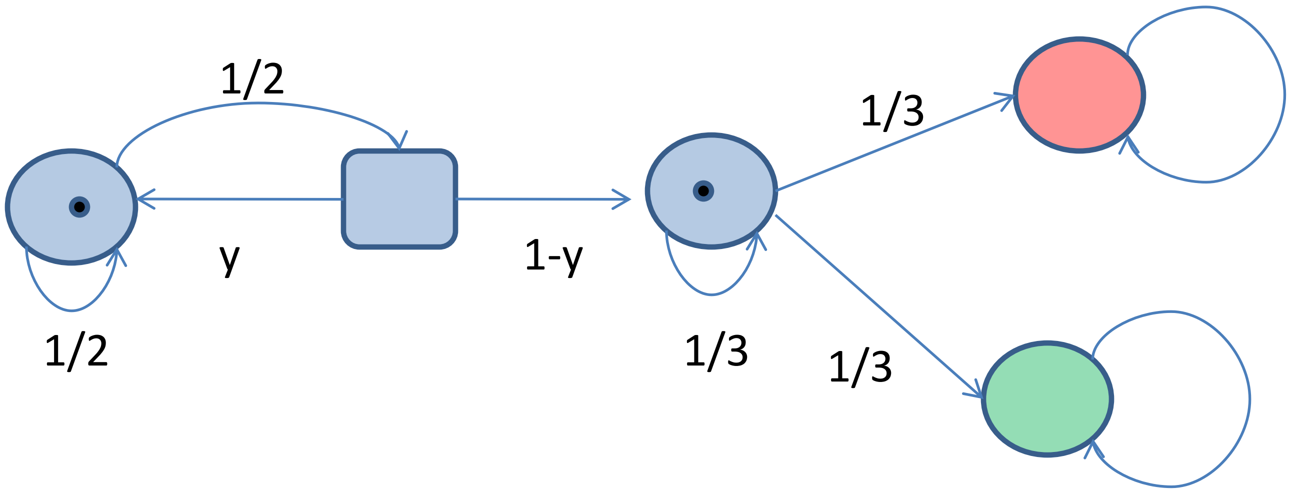
MDP: Example



MDP: Example



MDP: Example



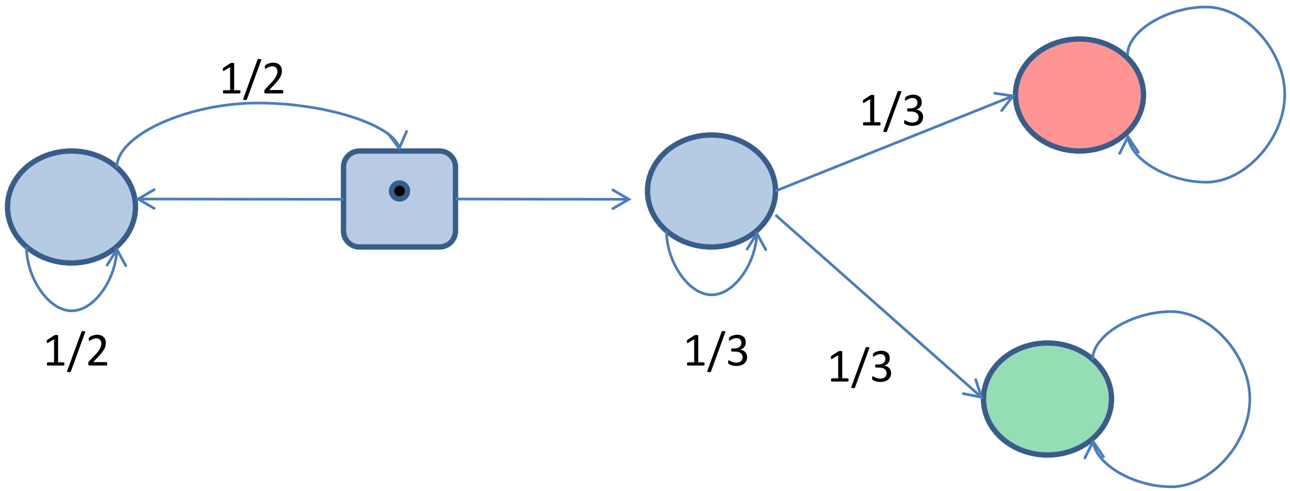
MDP

- 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: $S^* S_1 \rightarrow D(S)$.

MDP: Strategy Example

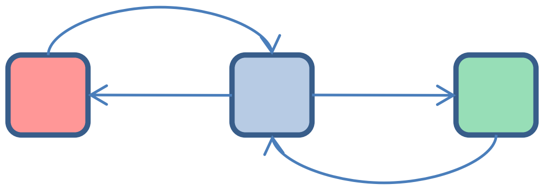


Token for k-th time: choose left with prob 1/k and right (1-1/k).

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
  - Val(ψ)(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.

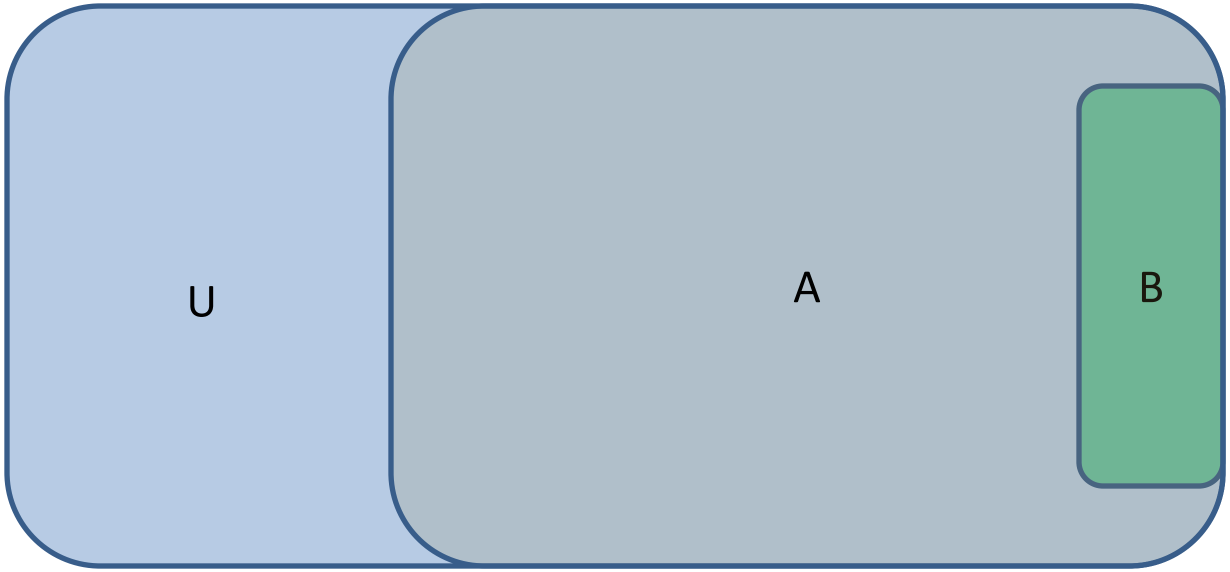
Attractor

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

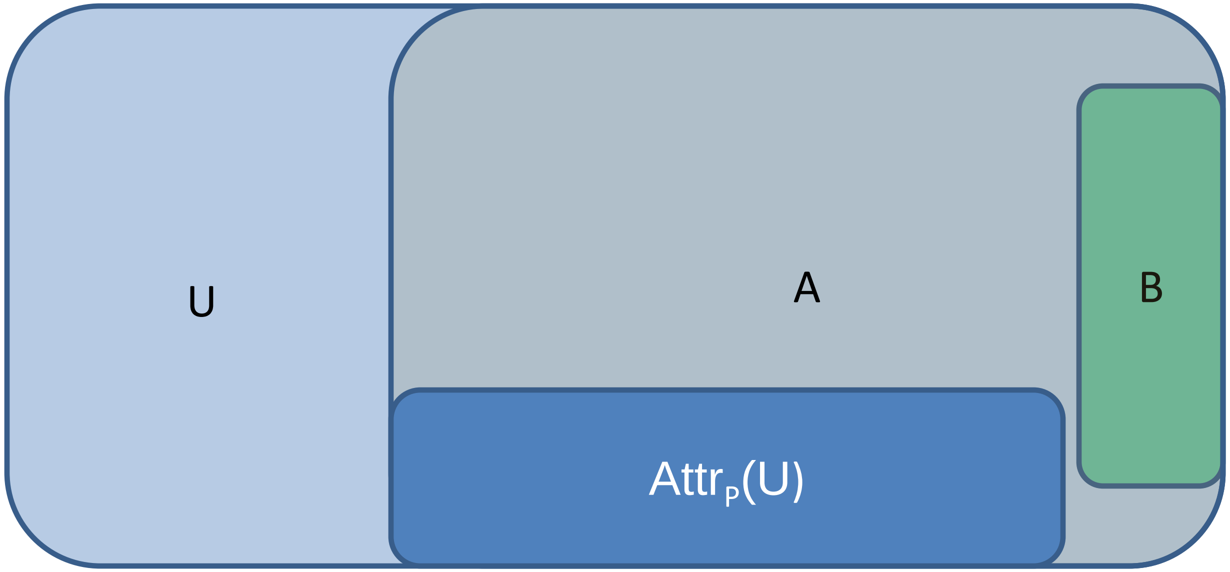
- Compute simple reachability to B (exist a path) in the graph of the MDP (S,E). Let us call this set A.

Iterative Algorithm



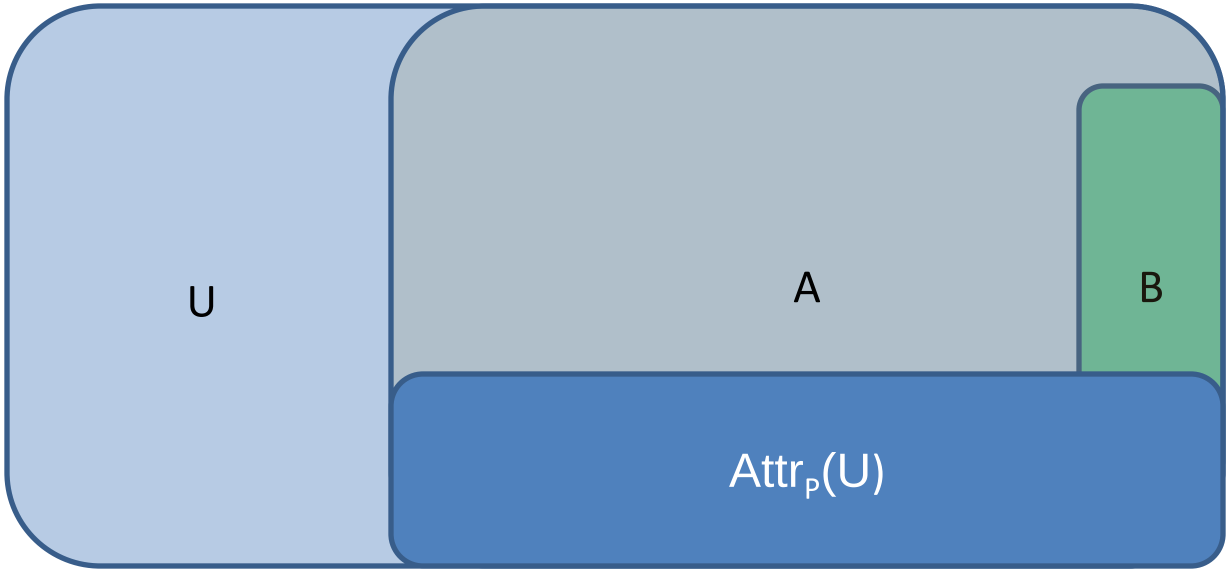
- Let U= S \ A. Then there is not even a path from U to B. Clearly, U is not in the almost-sure set.
- By attractor lemma can take $Attr_P(U)$ out and iterate.

Iterative Algorithm



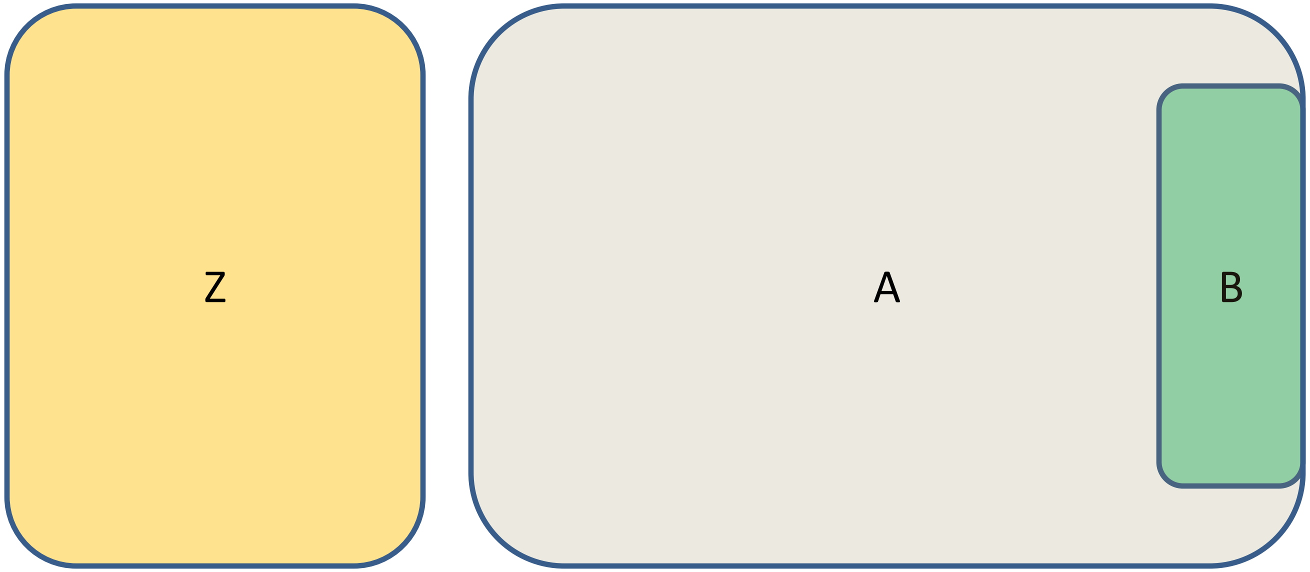
- $Attr_P(U)$ may or may not intersect with B.

Iterative Algorithm



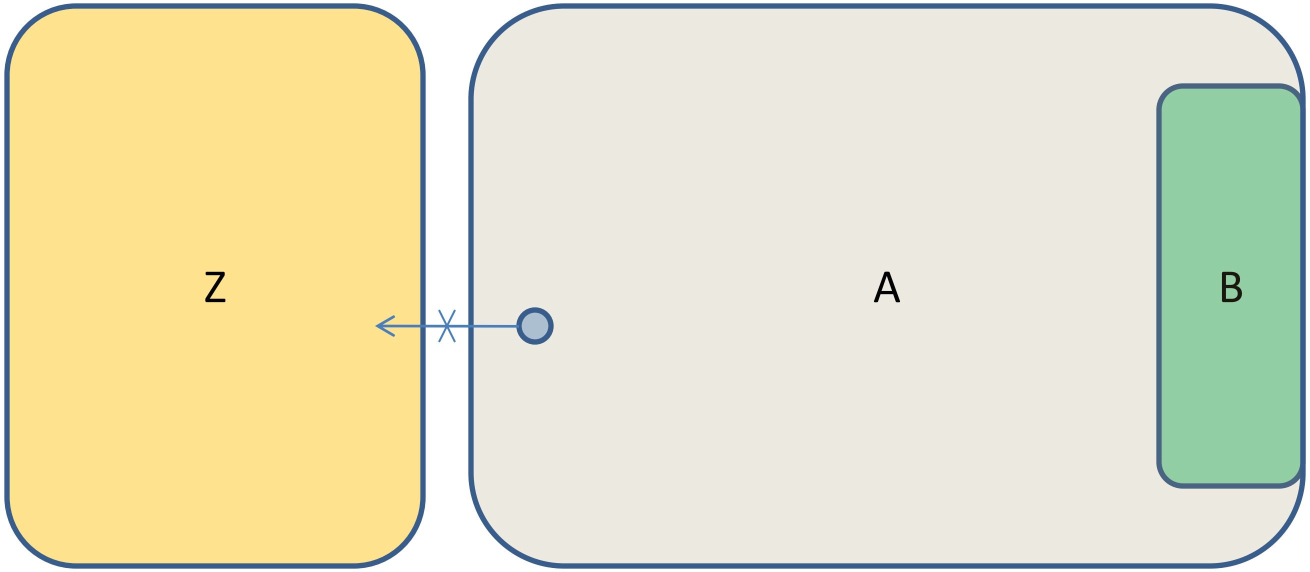
- Iterate on the remaining sub-graph.
- Every iteration what is removed is not part of almost-sure winning set.
- What happens when the iteration stops.

Iterative Algorithm



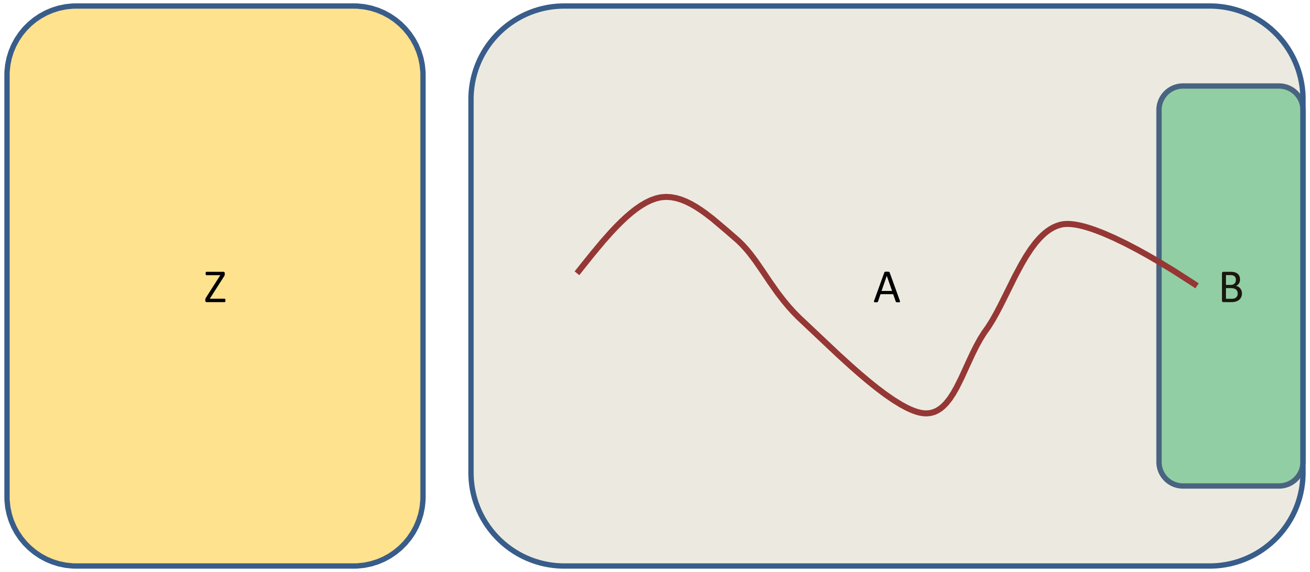
- The iteration stops. Let Z be the set of states removed overall iterations.
- Two key properties.

Iterative Algorithm



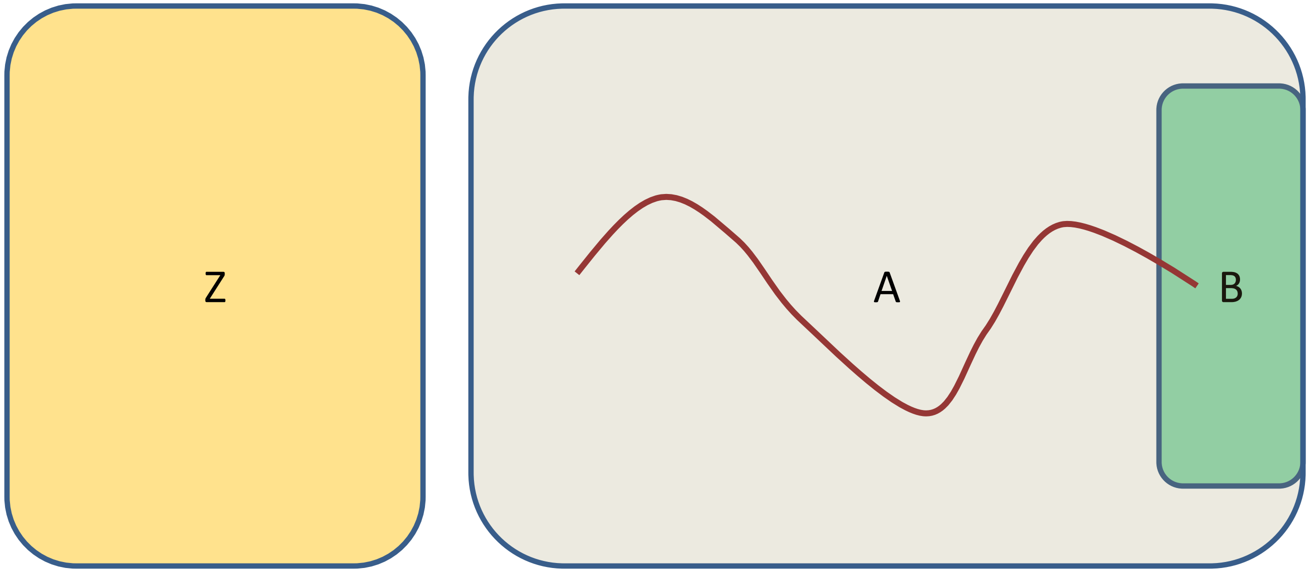
- The iteration stops. Let Z be the set of states removed overall iterations.
- Two key properties:
  - No probabilistic edge from outside to Z.

Iterative Algorithm



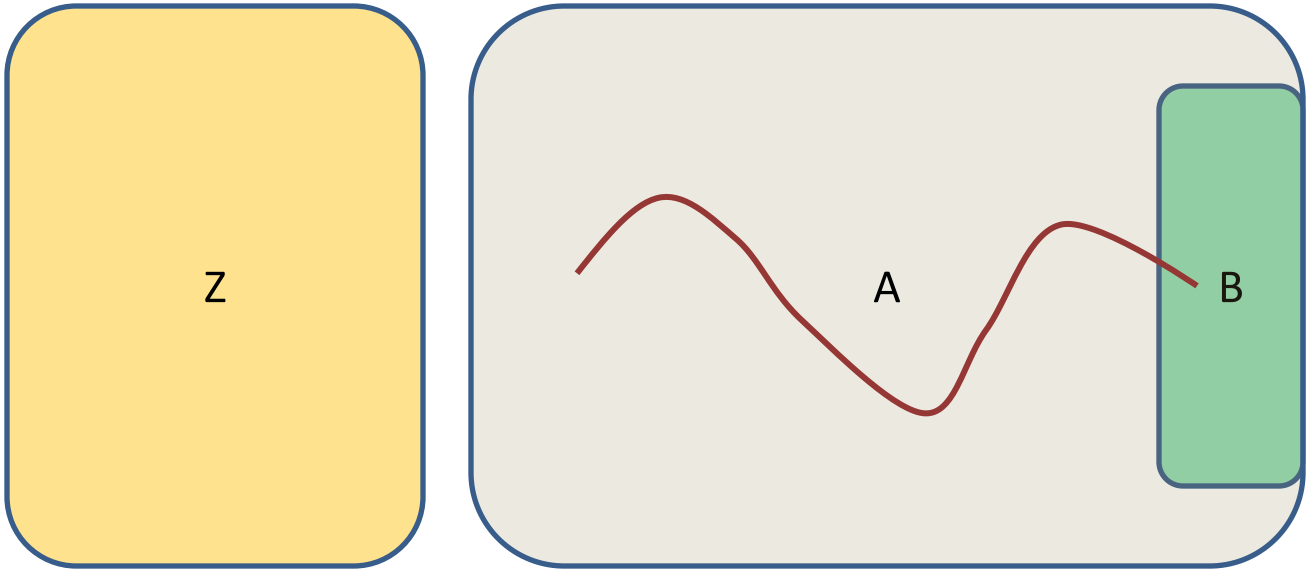
- The iteration stops. Let Z be the set of states removed overall iterations.
- Two key properties:
  - No probabilistic edge from outside to Z.
  - From everywhere in A (the remaining graph) path to B.

Iterative Algorithm



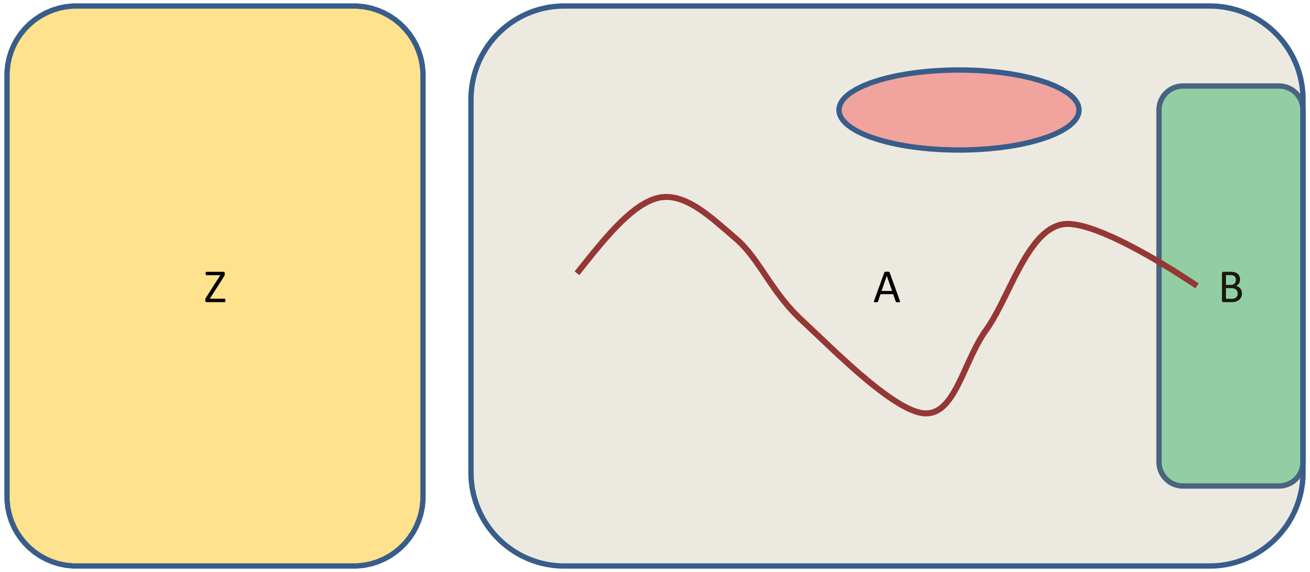
- Two key properties:
  - No probabilistic edge from outside to Z.
  - From everywhere in A (the remaining graph) path to B.
- Fix a memoryless strategy as follows:
  - In A $\setminus$ B: shorten distance to B. (Consider the BFS and choose edge).
  - In B: stay in A.

Iterative Algorithm



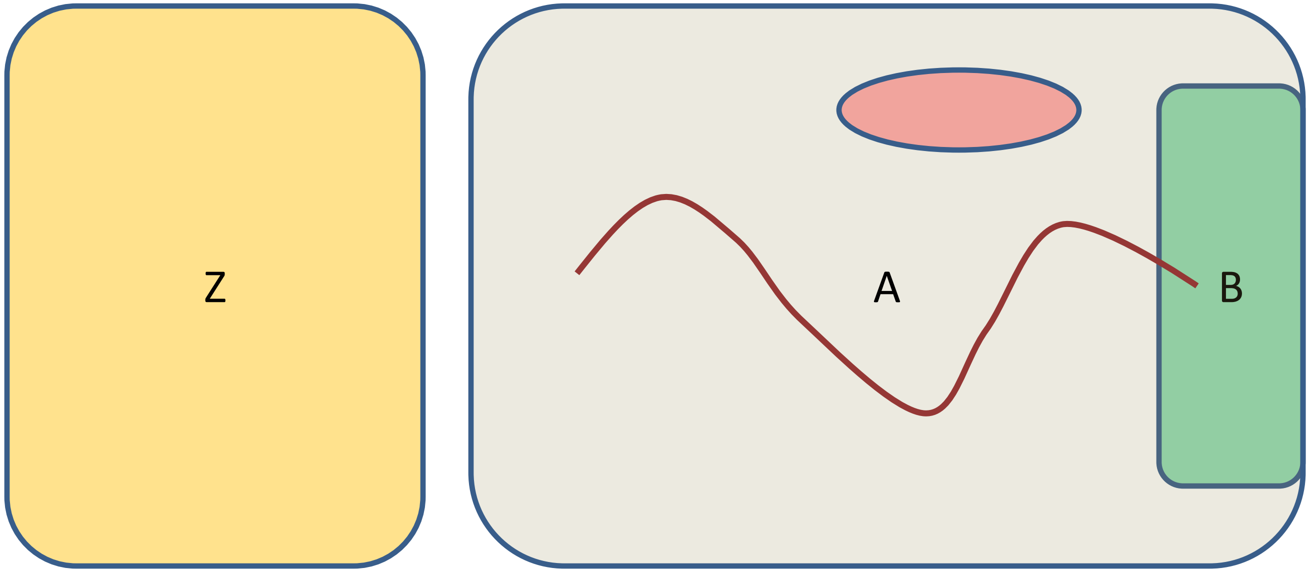
- Fix a memoryless strategy as follows:
  - In A \ B: shorten distance to B. (Consider the BFS and choose edge).
  - In B: stay in A.
- Argue all bottom scc's intersect with B. By Markov chain theorem done.

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.

Quantitative Parity to Quantitative Reachability

- End-components: An end-component generalizes both scc and closed recurrent set. A set U is an end-component if the following properties hold:
  - U is strongly connected.
  - U is closed (no probabilistic edge out).

- Note that player 1 edges may leave the end-component.

- Why is end-component important: it allows us to reason about infinite behaviors.

End Component Property

- End-component property: For an MDP and for all strategies, with probability 1 the set of states visited infinitely often is an end-component.

- Generalizes the scc for graphs and closed recurrent set for Markov chains.

- Proof:
  - Shape of the proof very similar to closed recurrent set.
  - We need to show that if a set U is not an end-component, then cannot be visited infinitely often with positive probability.
  - Assume towards contradiction that there is such a set U.

End Component Property

- End-component property: For an MDP and for all strategies, with probability 1 the set of states visited infinitely often is an end-component.
- Proof:
  - We need to show that if a set U is not an end-component, then cannot be visited infinitely often with positive probability.
  - Assume towards contradiction that there is such a set U.
  - U must be strongly connected.
  - Since U is not end-component, some probabilistic state s with an edge to t going out of U with probability α.
  - Hence the probability that s is visited infinitely often, but the edge to t is taken finitely often is 0.
  - The result follows.

Winning End-component

- An end-component U is winning if the minimum priority of U is even.

- From end-component property for any strategy the probability to satisfy parity is the probability to reach the winning end-components.

- In winning end-components pure memoryless almost-sure winning strategy exists.
    - Proof: Choose successor to shorten distance to the minimum even priority state.

Quantitative Parity to Quantitative Reachability

- The probability to satisfy is the probability to reach winning end-components.
- In winning end-components pure memoryless almost-sure strategy.
- Winning end-components are included in the almost-sure winning set.
- Hence we need quantitative reachability to almost-sure winning set.
- We now need the quantitative reachability to complete the argument.

Quantitative Reachability

- An MDP G, and a target set T.
- Val(Reach(T))(s) = $\sup_\sigma \Pr_s^\sigma$(Reach(T)).

- v(s) for abbreviation.

- Two properties:
  - Property 1: For $s \in S_P$ we have $v(s) = \sum_{t \in S} v(t) * \delta(s)(t)$.
  - Property 2: For $s \in S_1$ we have $v(s) = \max \{ v(t) \mid t \in E(s) \}$.

Proof of Property 2

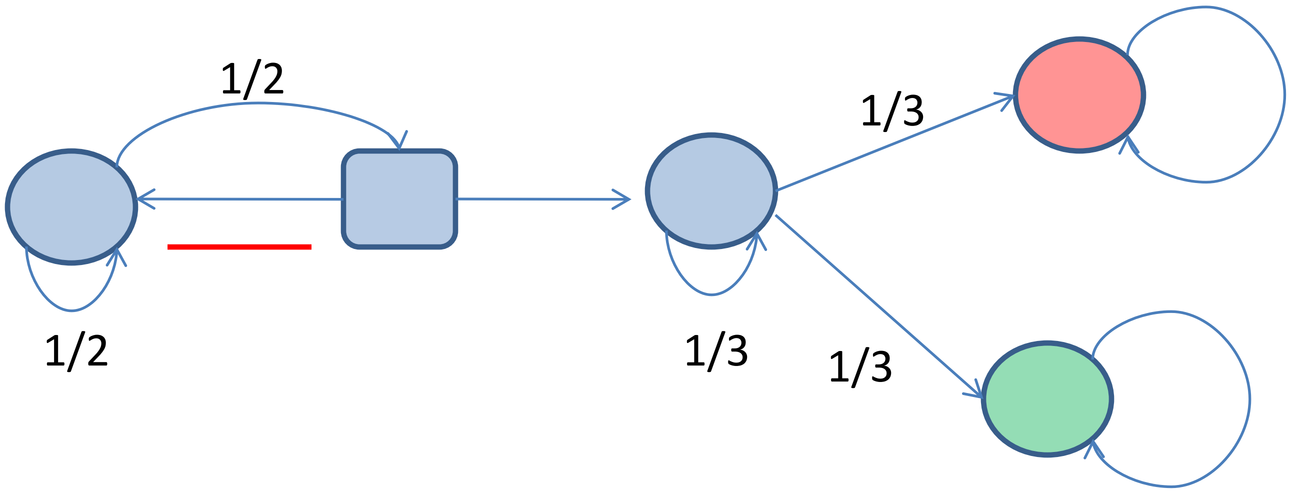
- Inequality 1: $v(s) \geq \max\{ v(t) \mid t \in E(s)\}$
  - Fix ϵ>0.
  - Let t* be the arg max.
  - From s choose t*, and then an ϵ optimal strategy from t* to ensure value at least v(t*)-ϵ.
  - As ϵ>0 is arbitrary, the result follows.

- Inequality 2. $v(s) \leq \max\{ v(t) \mid t \in E(s)\}$
  - We have
    - $v(s) \leq \sup_{\mu} \sum_{t \in S} v(t) * \mu(t) \leq \max \{v(t) \mid t \in E(s)\}$, where $\mu \in D(E(s))$.

A Simple Attempt

- For a state s choose a successor that achieves the maximum.
- However this simple construction is not sufficient.

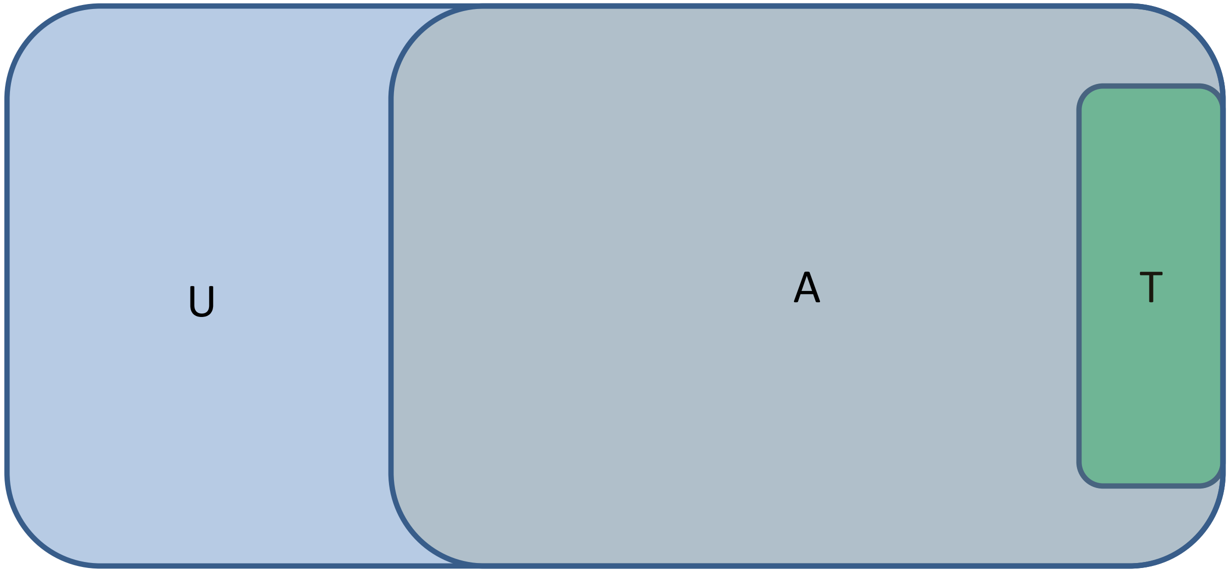
MDP: Simple Fails



In all blue states the value is ½.

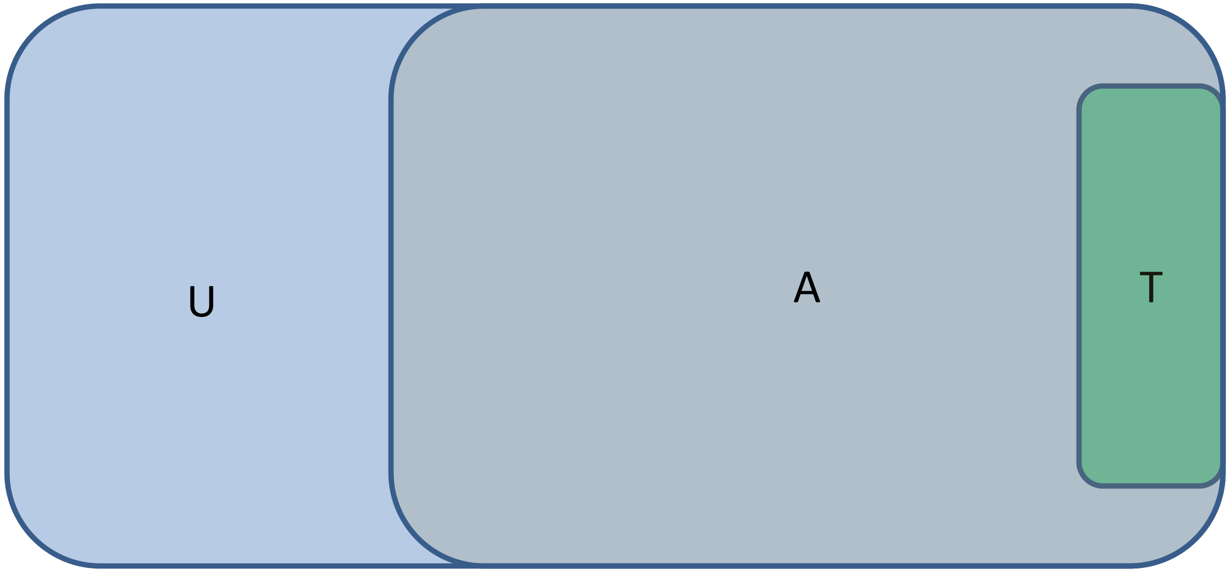
However the choice of red edge is bad.

Quantitative Reachability



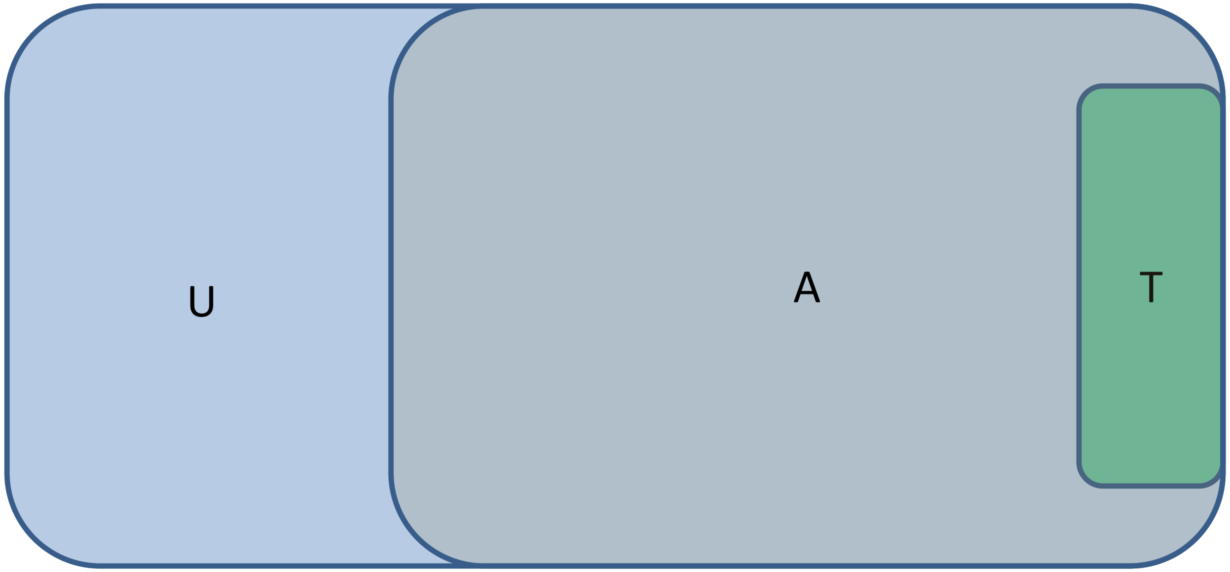
- Original MDP is connected.
- Compute simple reachability to T.
- From U, there is no path so value is 0.
- From A, the value is positive everywhere as there is a path.

Quantitative Reachability



- From U, there is no path so value is 0.
- From A, the value is positive everywhere as there is a path.
- Retain only the edges that attains the max in A (remove all the other). Make U and T absorbing.
- Easy to show that there is still path to T from A.
- Choose the edge that shortens distance to T.

Quantitative Reachability



- Retain only the edges that attains the max in A (remove all the other). Make U and T absorbing.
- Easy to show that there is still path to T from A.
- Choose the edge that shortens distance to T.
- Markov chain where all closed recurrent states are U or T.
- The values v(s) satisfies the Markov chain equality. Hence the memoryless strategy achieves v(s).

Quantitative Reachability

- An MDP G, and a target set T.
- $\mathrm{Val}(\mathrm{Reach}(T))(s) = \sup_{\sigma} \Pr_s^{\sigma}(\mathrm{Reach}(T))$.

- Existence of pure memoryless optimal strategies.

- Algorithm: Linear programming. Variable x_s for all states s.
  - $x_s = 0$ $\quad s \in U$
  - $x_s = 1$ $\quad s \in T$
  - $x_s = \sum_{t \in S} x_t * \delta(s)(t)$ $\quad s \in S_P$
  - $x_s = \max_{t \in E(s)} x_t$ $\quad s \in S_1$.

Quantitative Reachability

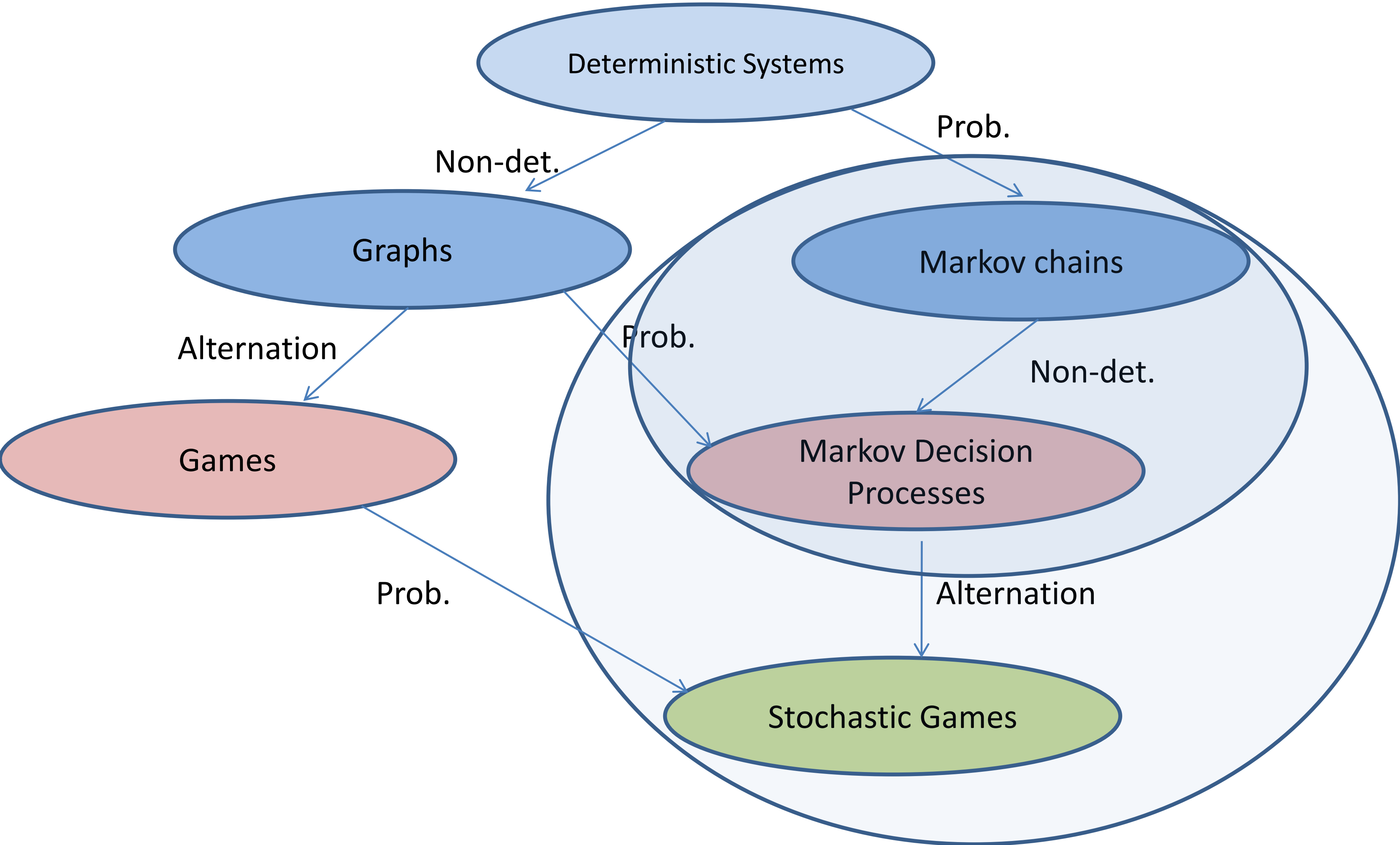
- Algorithm: Linear programming. Variable x_s for all states s.
  - $x_s = 0$ $\quad s \in U$
  - $x_s = 1$ $\quad s \in T$
  - $x_s = \sum_{t \in S} x_t * \delta(s)(t)$ $\quad s \in S_P$
  - $x_s = \max_{t \in E(s)} x_t$ $\quad s \in S_1$.

- The above optimization to linear program
  - Objective function: min $\sum_{t \in S} x_t$
  - $x_s \geq x_t$ $\quad s \in S_1, t \in E(s)$.

MDP Summary

	Reachability	Liveness	Parity
Qualitative	O(n m)	O(n m)	O(n m d)
Quantitative	Linear programming	Linear programming	Linear programming

Stochastic Games

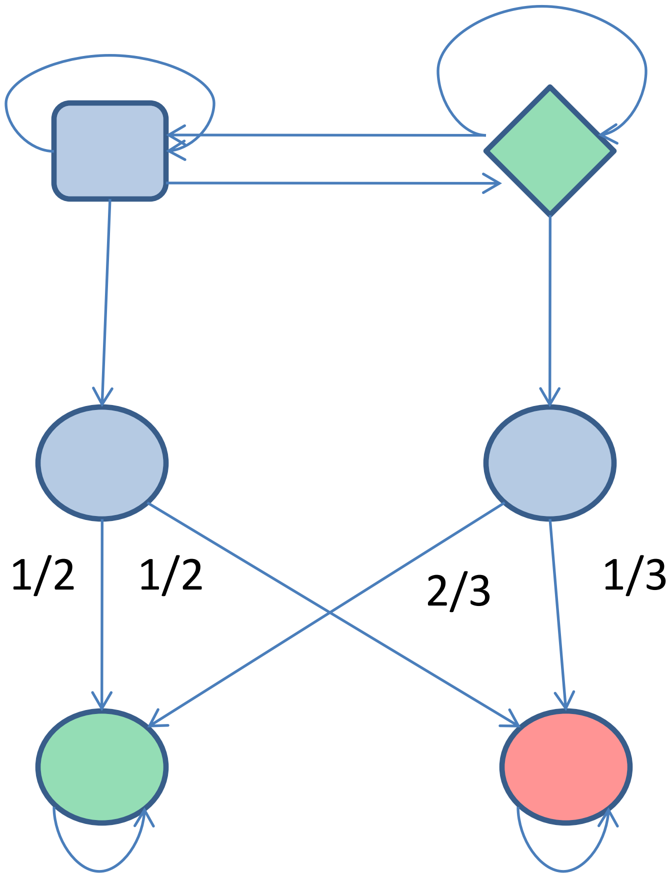


Stochastic Games

- Stochastic games
  - Non-determinism: angelic vs. demonic non-determinism (alternation).
  - Probability.
  - Generalizes non-deterministic systems and Markov chains, alternating games, MDPs.

- An MDP G= ((S,E), (S_1, S_2,S_P), δ)
  - $\delta : S_P \rightarrow D(S)$.
  - 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.

Stochastic Game



Example of stochastic game.

Objective for player 1 is to visit green infinitely often

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: $S^* S_1 \rightarrow D(S)$.
  - π: $S^* S_2 \rightarrow D(S)$.

Strategies

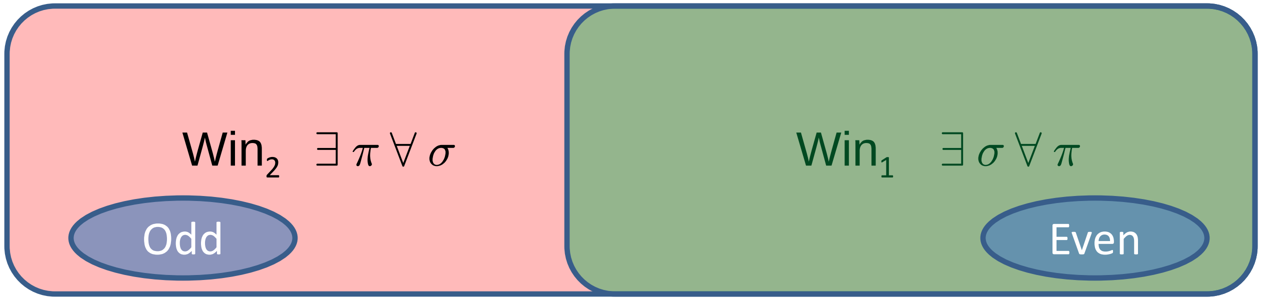
- Strategies are recipe how to move tokens or how to extend plays. Formally, given a history of play (or finite sequence of states), it chooses a probability distribution over out-going edges.
  - σ: $S^* S_1 \to D(S)$.
- History dependent and randomized.
- History independent: depends only current state (memoryless or positional).
  - σ: $S_1 \to D(S)$
- Deterministic: no randomization (pure strategies).
  - σ: $S^* S_1 \to S$
- Deterministic and memoryless: no memory and no randomization (pure and memoryless and is the simplest class).
  - σ: $S_1 \to S$
- Same notations for player 2 strategies π.

Values in Stochastic Games

- Value at a state for an objective ψ
  - $\mathrm{Val}(\psi)(s) = \sup_{\sigma} \inf_{\pi} \Pr_s^{\sigma,\pi}(\psi)$.
- 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.
- Determinacy: the order of sup inf can be exchanged.

Non-Stochastic Games

- There are no probabilistic states.
- Non-stochastic games with parity objectives
  - Values only 0 or 1.
  - Pure memoryless winning strategies exist.
  - Once a pure memoryless strategy is fixed all cycles winning.

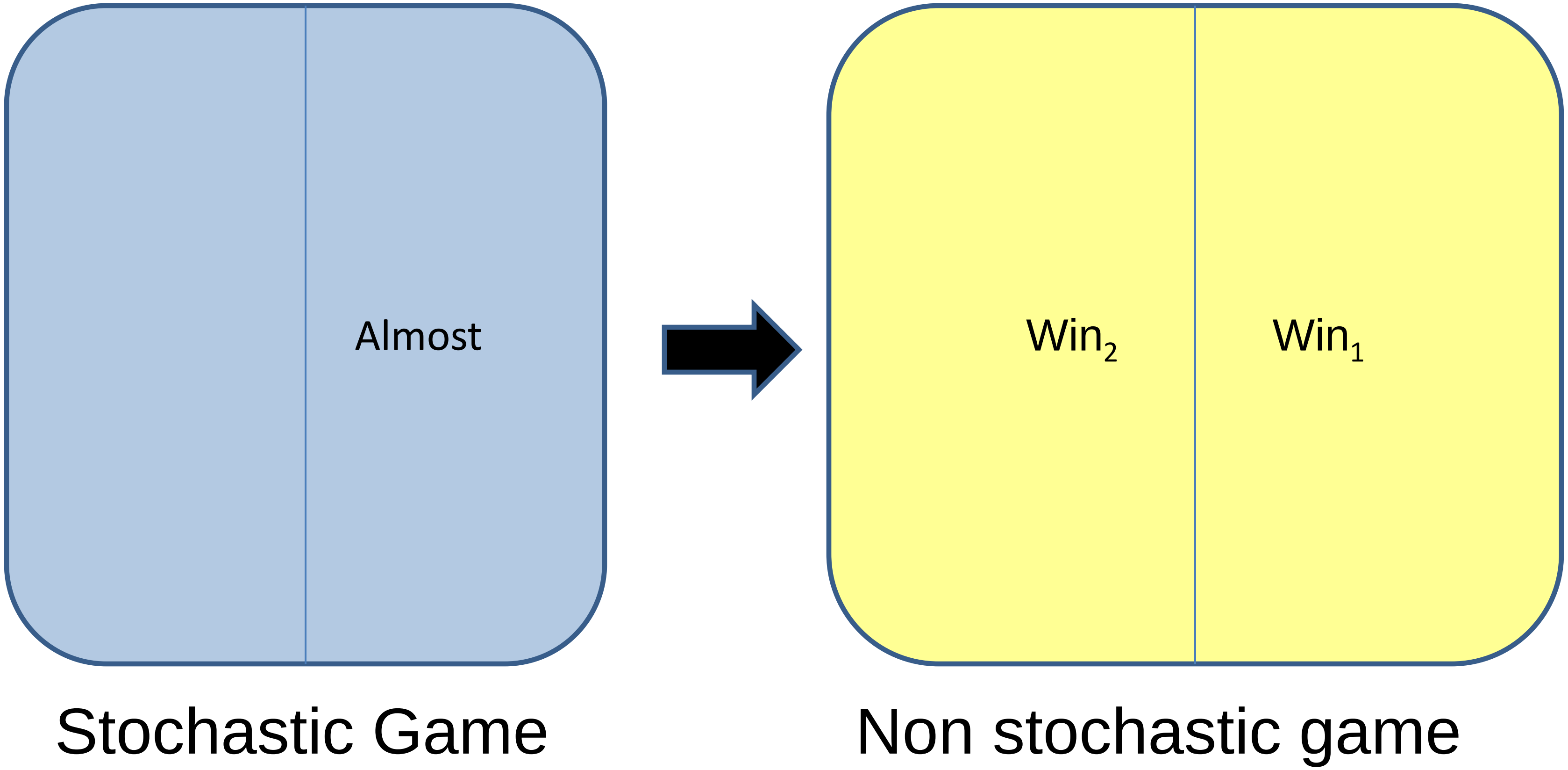


Qualitative and Quantitative Analysis

- Qualitative analysis
  - Reduction to games without probability.
  - Use existence of pure memoryless strategies in games with probability for parity objectives.
  - Show it for Liveness and can be extended to parity.
- Quantitative analysis
  - Combine notion of qualitative and local optimality for quantitative optimality.

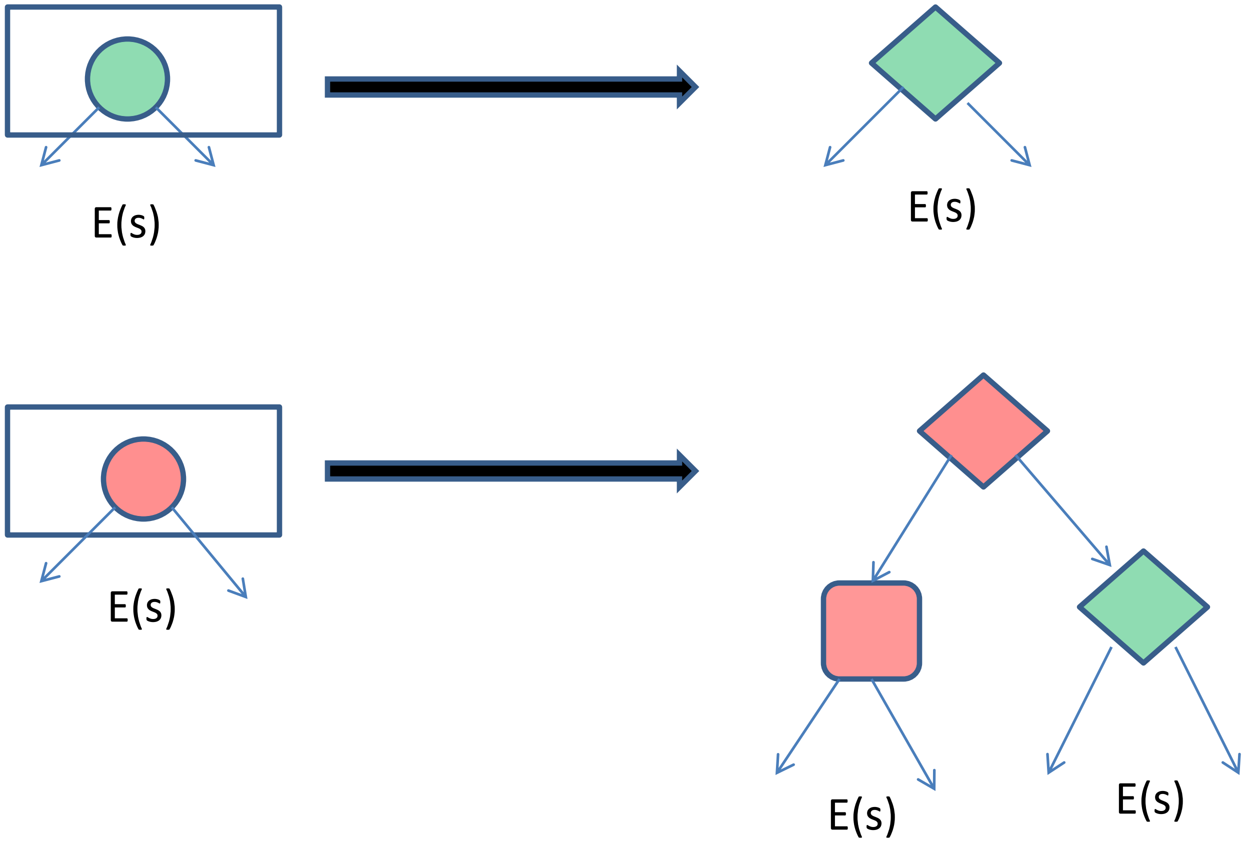
Qualitative Analysis

- Reduction



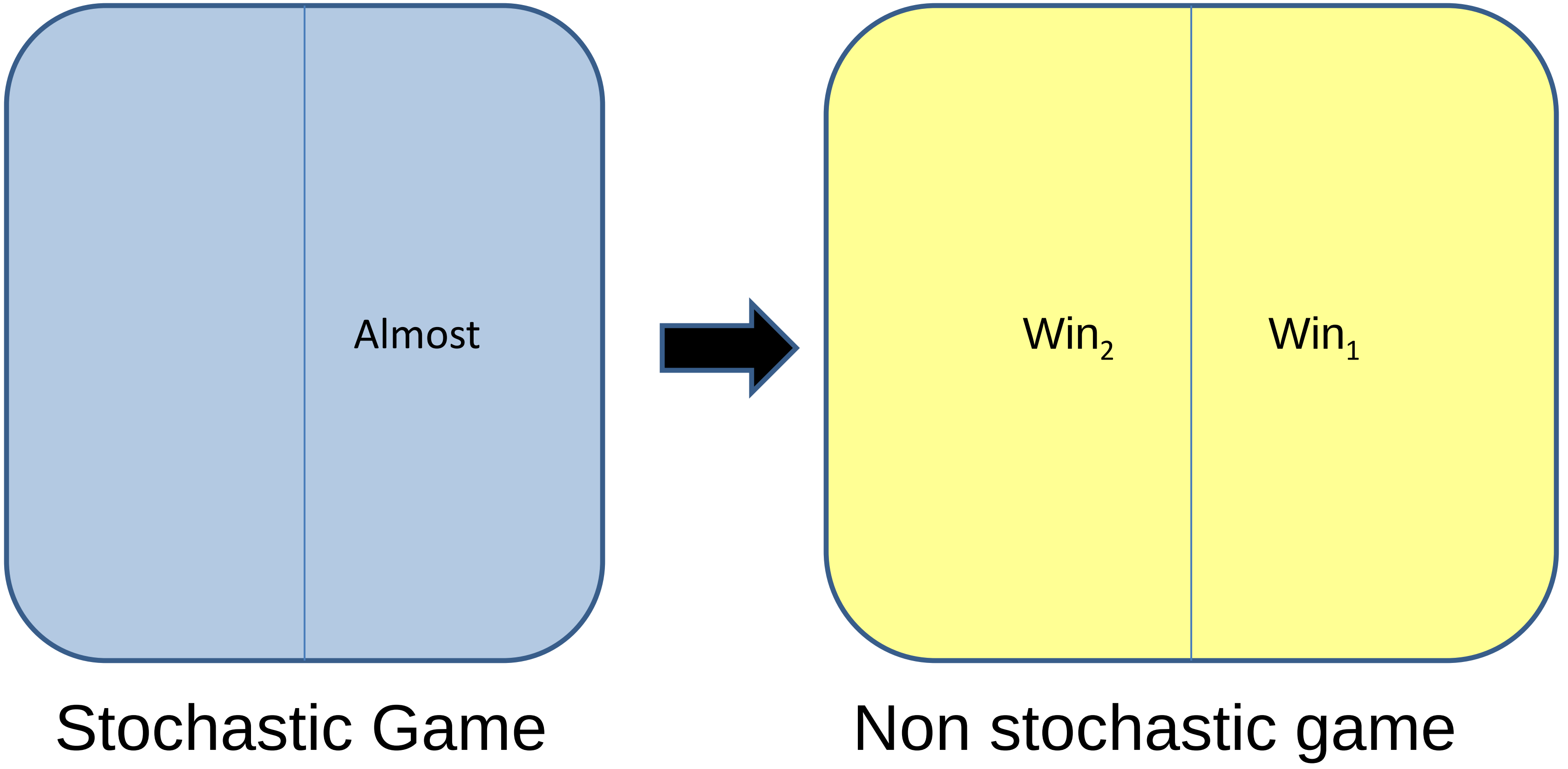
Reduction

- Replace every probabilistic state by two-player gadget. Illustrate it for Liveness.



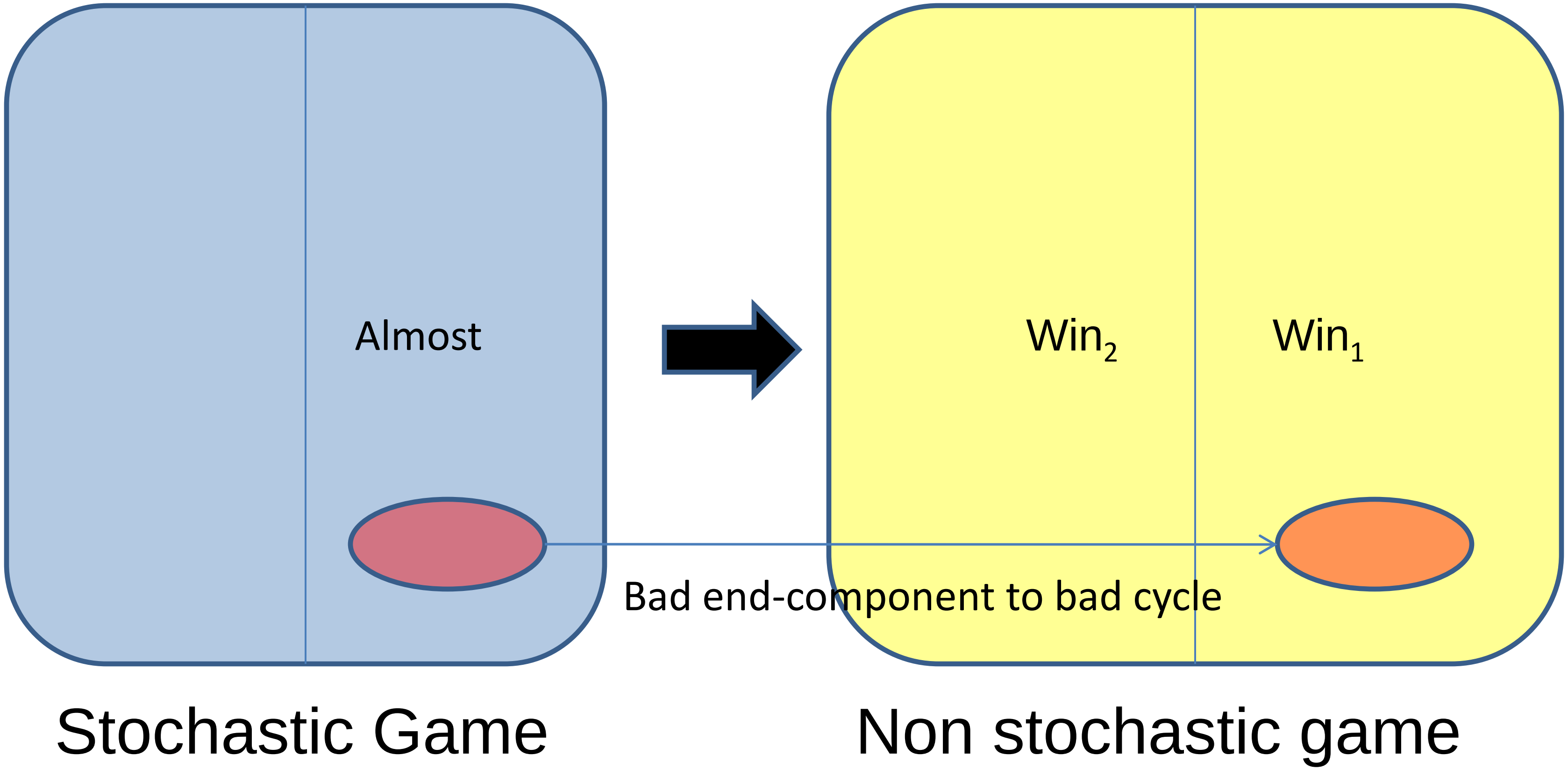
Qualitative Analysis

- Reduction: the end-components are winning.



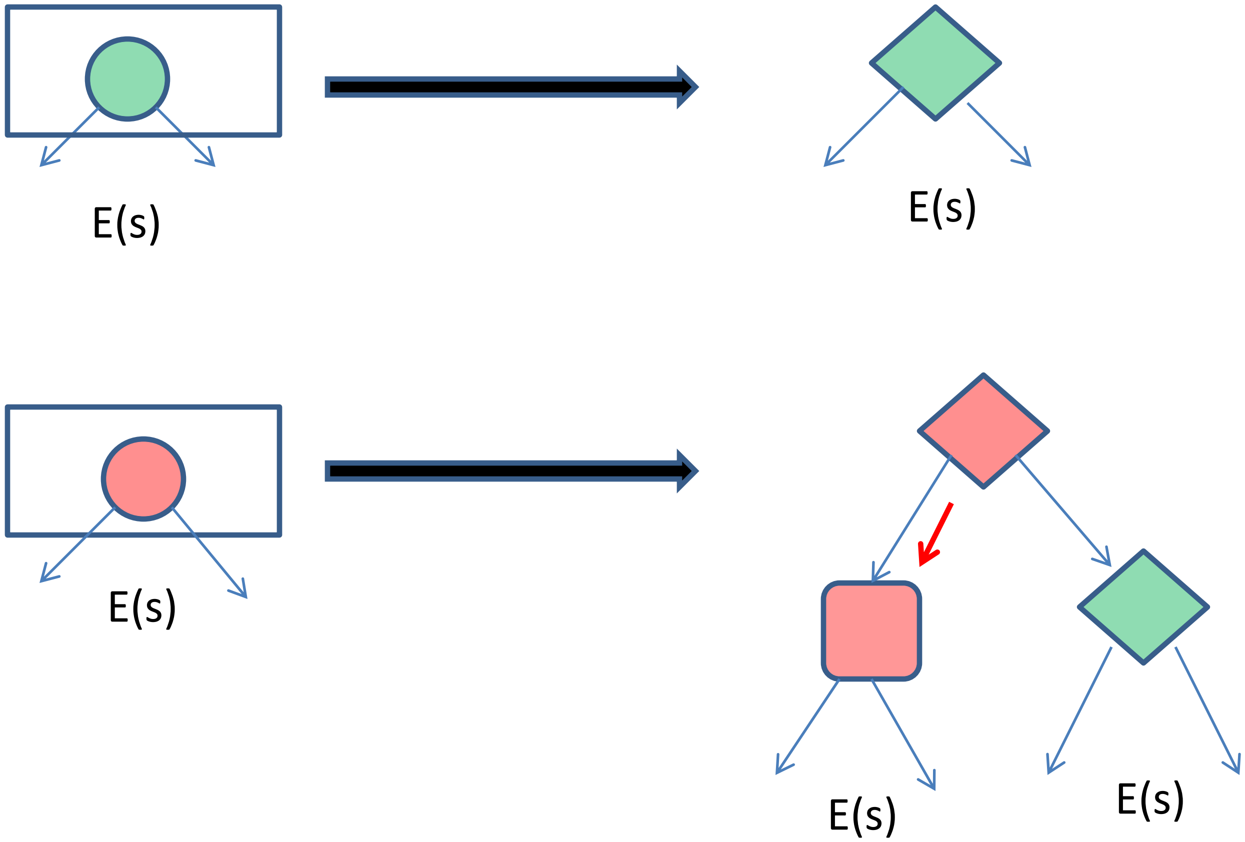
Qualitative Analysis

- Reduction: the end-components are winning.



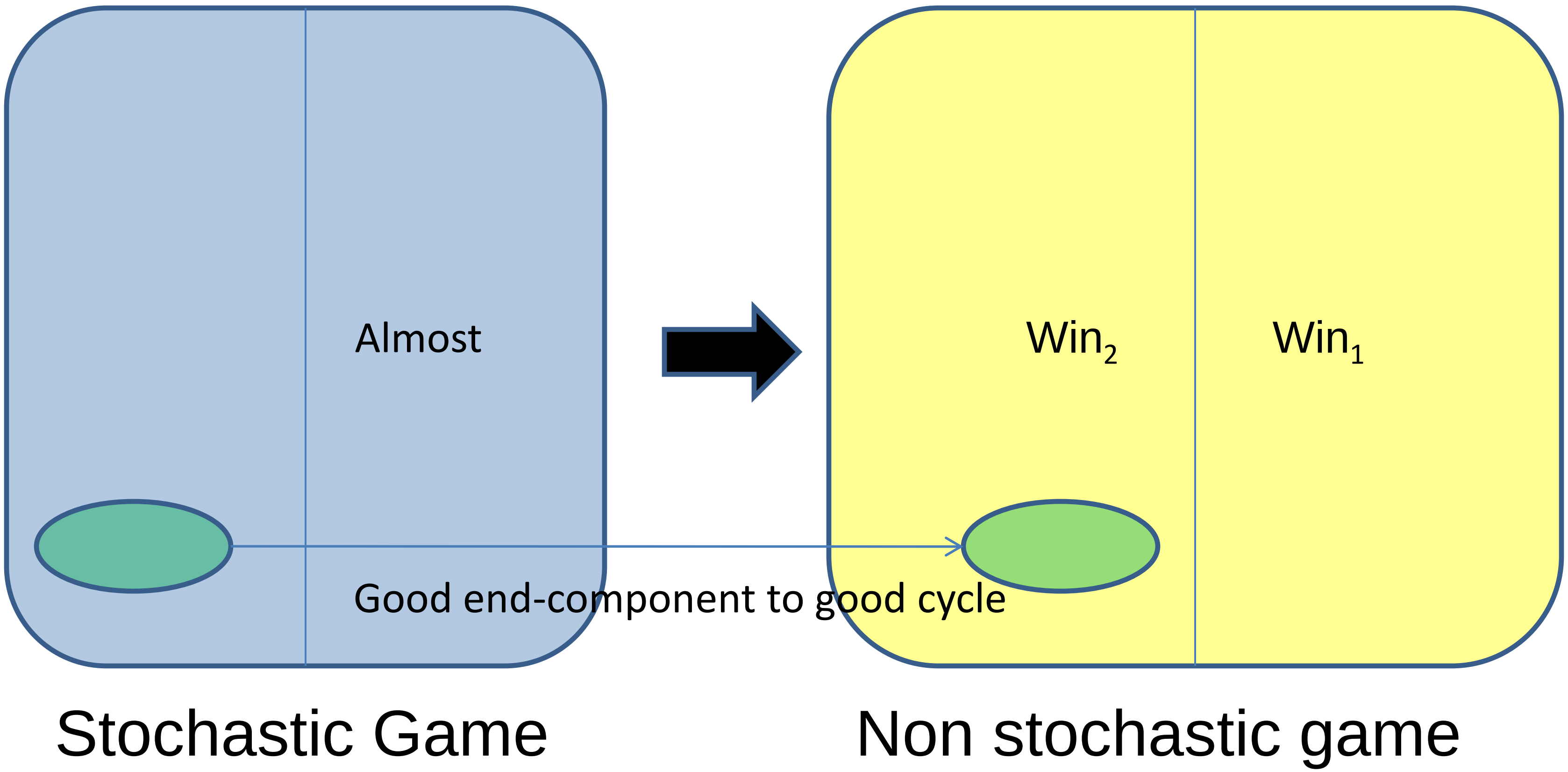
Reduction

- Choice in the gadget



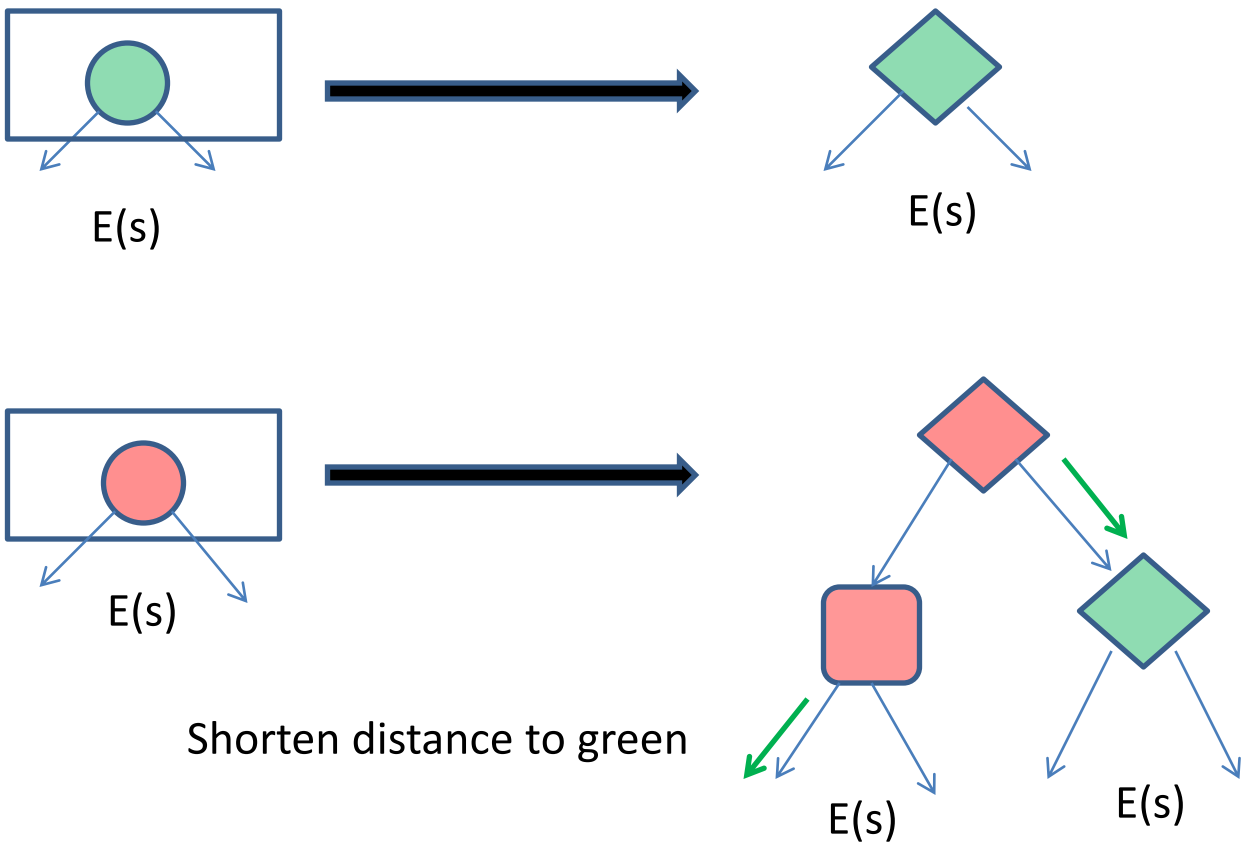
Qualitative Analysis

- Reduction: the end-components are winning.



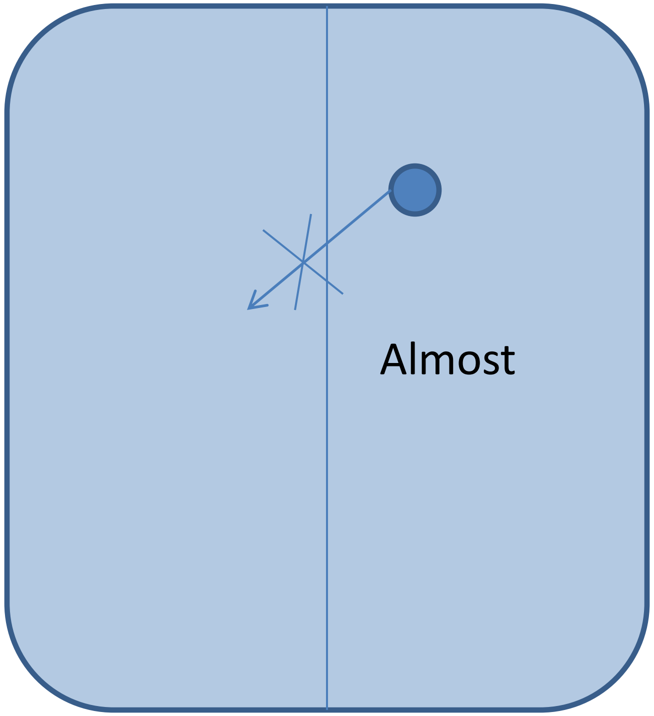
Reduction

- Choice in the gadget

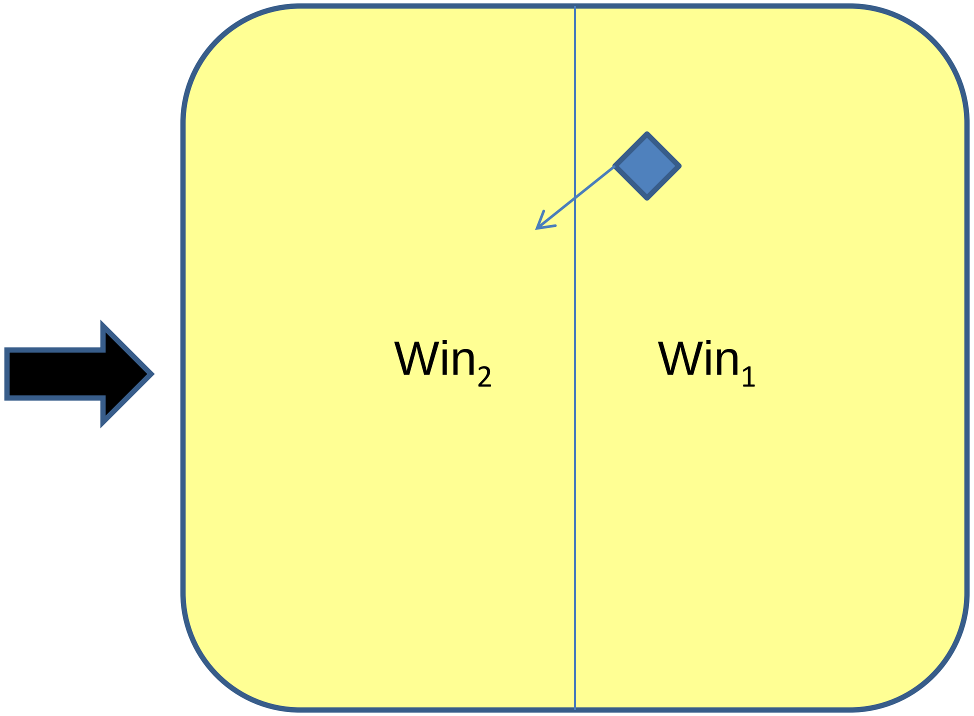


Qualitative Analysis

- Reduction: the end-components are winning.



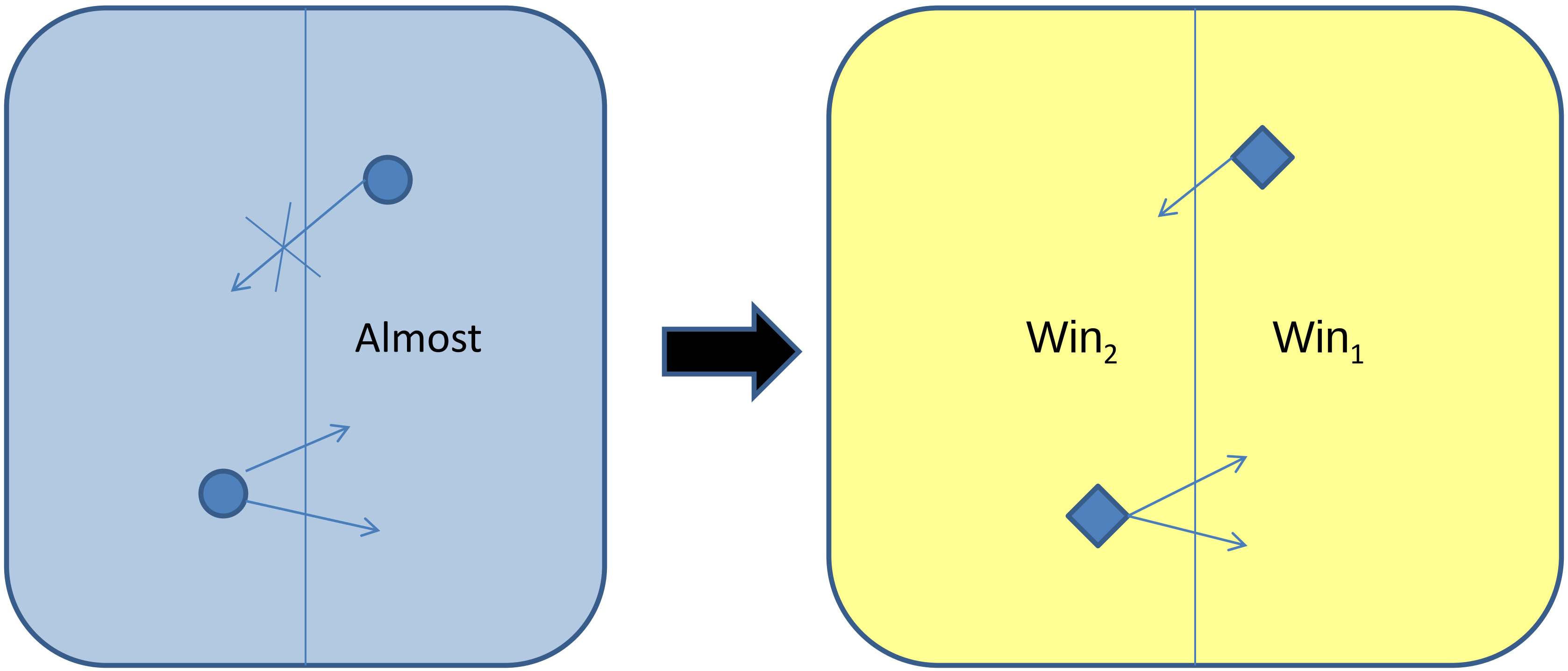
Stochastic Game



Non stochastic game

Qualitative Analysis

- Reduction: the end-components are winning.



Stochastic Game

Non stochastic game

Reduction

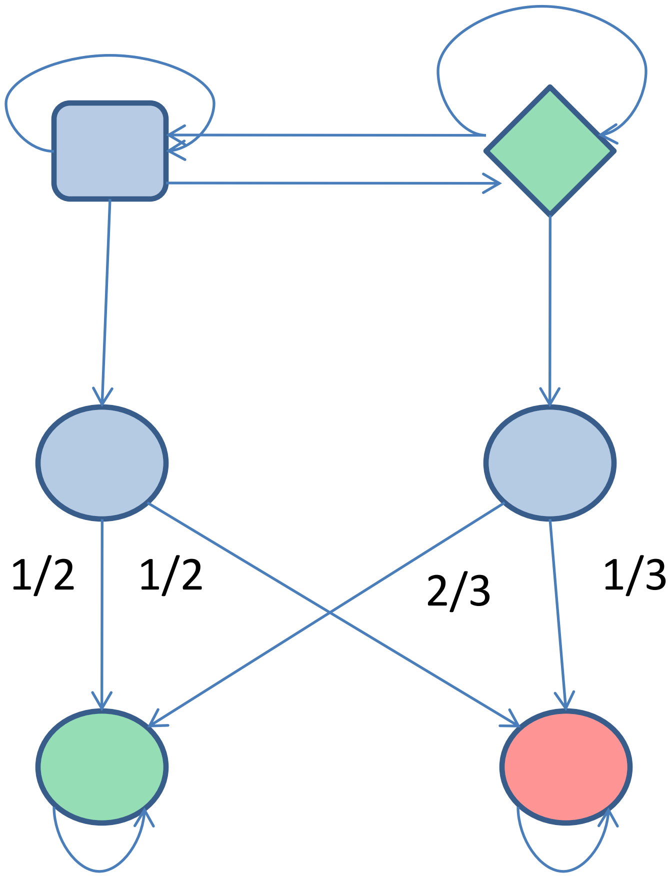
- Gadget based reduction can be extended to parity.

- Qualitative analysis
  - Pure memoryless almost-sure strategies exists.
  - Linear time reduction to non-stochastic games.
  - Same complexity: NP $\cap$ coNP.
  - All algorithms can be used.

Quantitative Analysis

- Unlike MDPs, we cannot do the following:
  - Compute almost-sure winning states.
  - Compute quantitative reachability to almost-sure winning states.
  - We illustrate with an example.

Stochastic Game



Example of stochastic game.

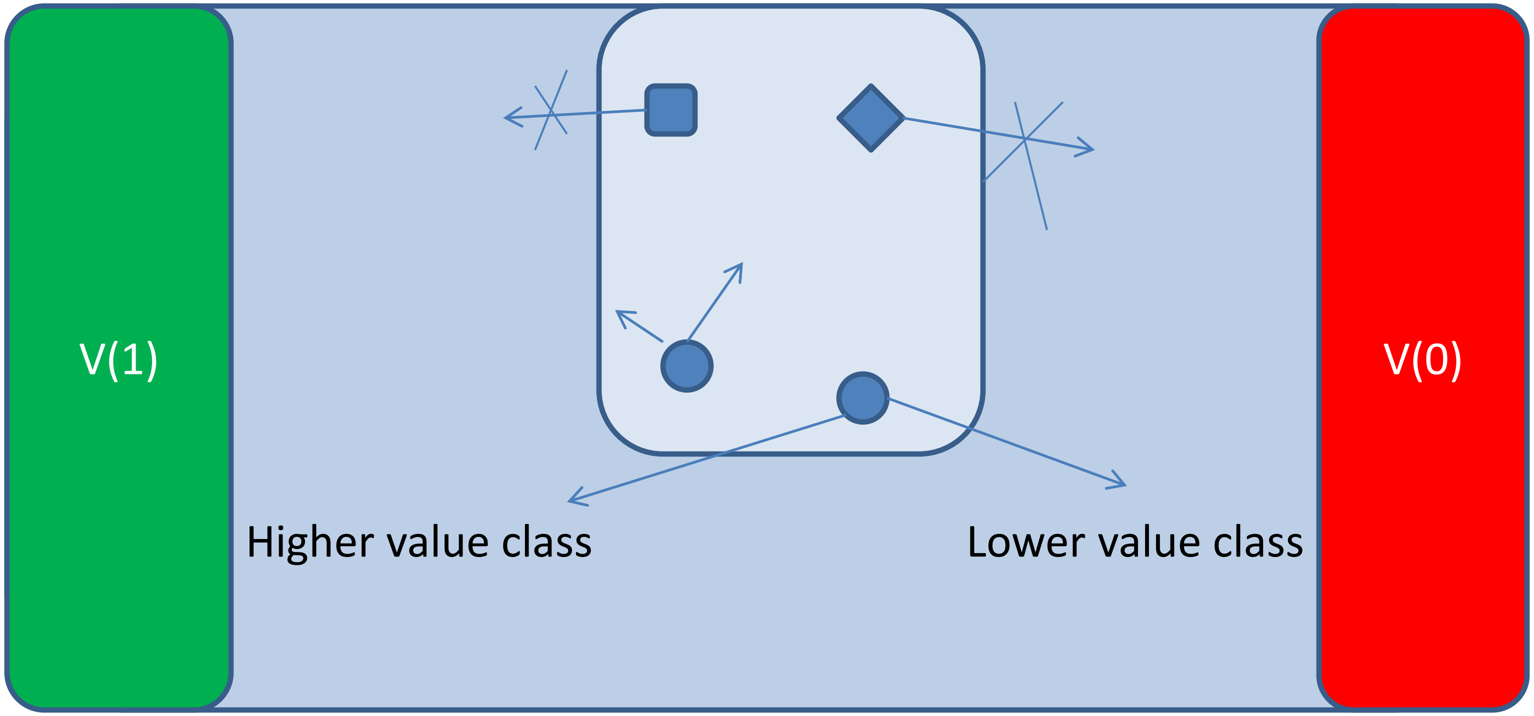
Objective for player 1 is to visit green infinitely often

Cannot ensure to reach green absorbing with prob 2/3.

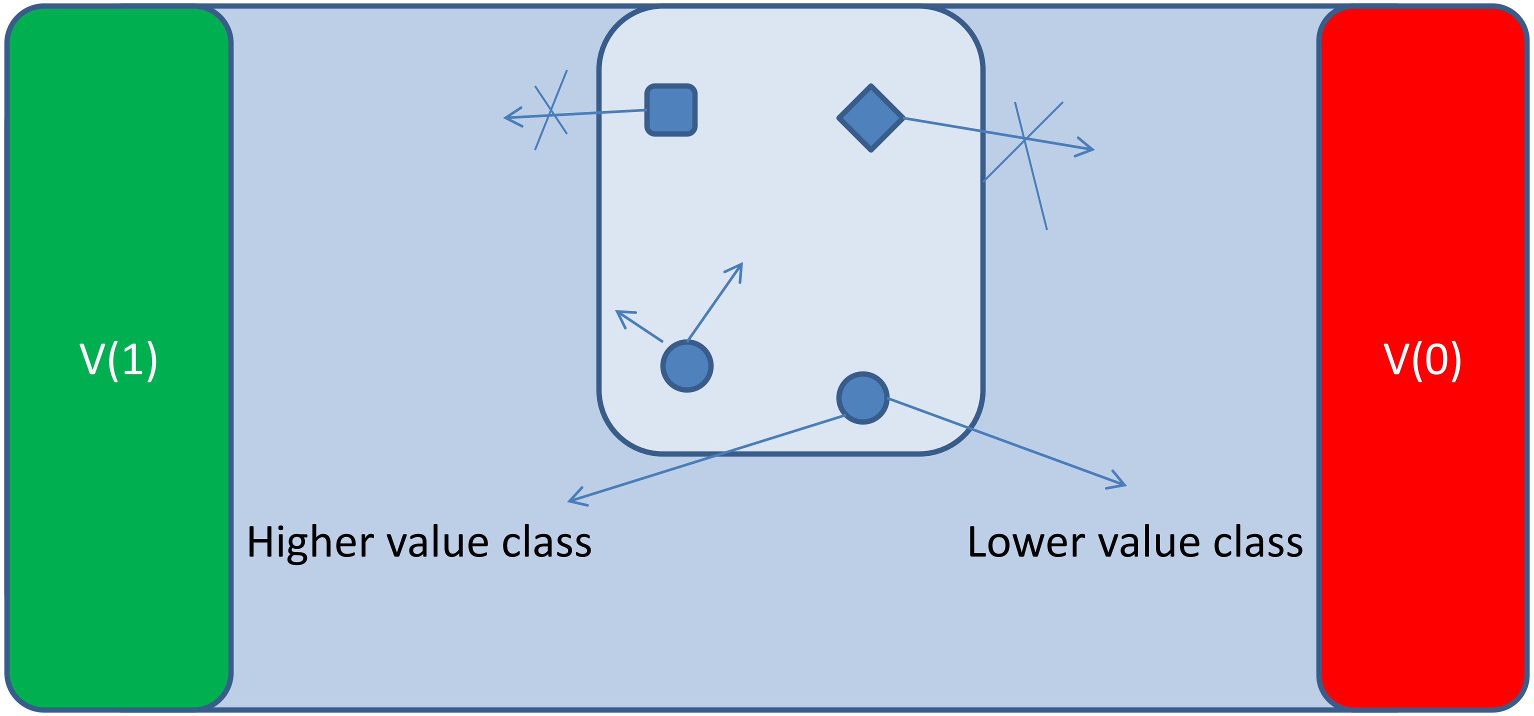
Quantitative Analysis

- Quantitative optimality
  - Local optimality
  - Qualitative optimality
- Value class: the set of states with same value. V(r) is the set of states with value r.

Value Class Property



Value Class: Boundary Probabilistic States



Value Class Reduction

- Remove edges going out to lower value class (local optimality).

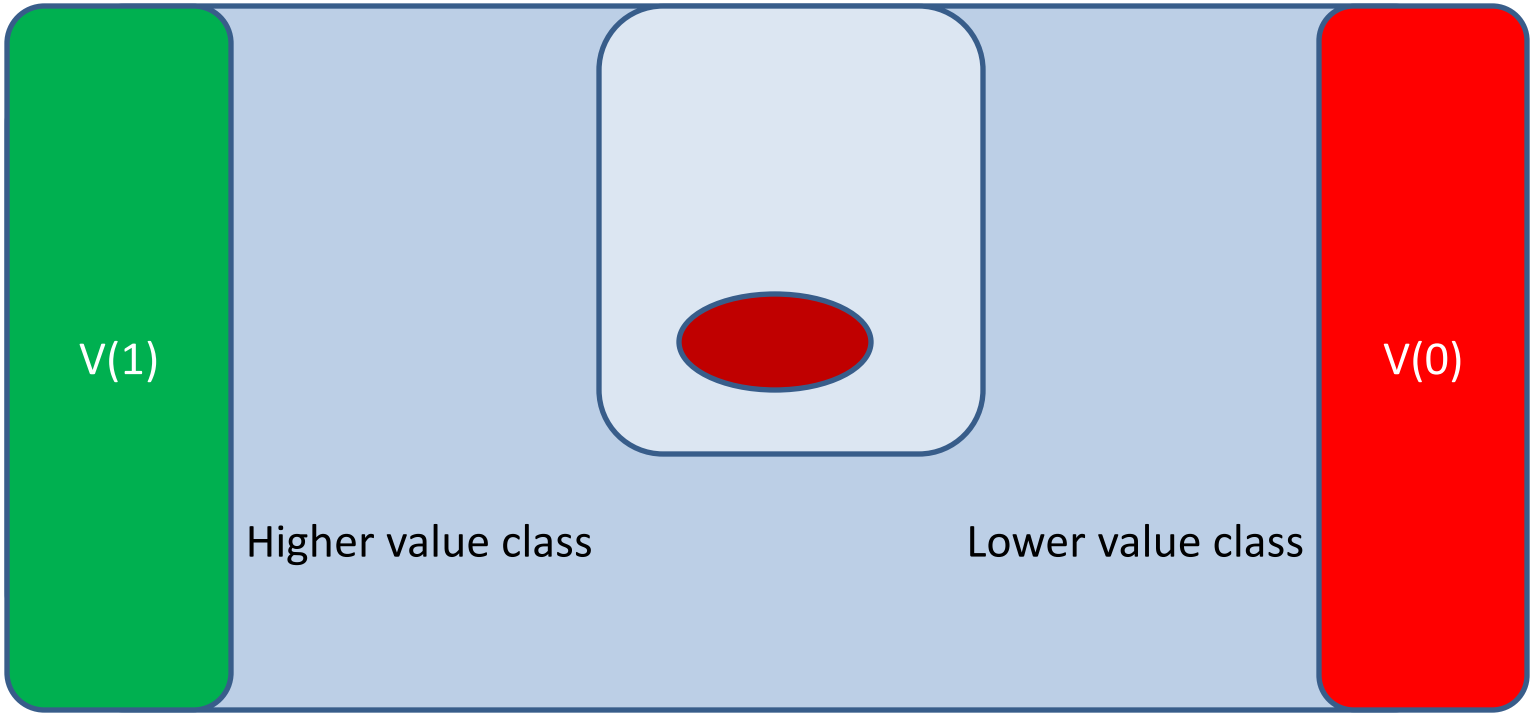
- Change boundary probabilistic states to winning states for player 1.

- Claim: In this sub-game player 1 wins almost-surely everywhere.

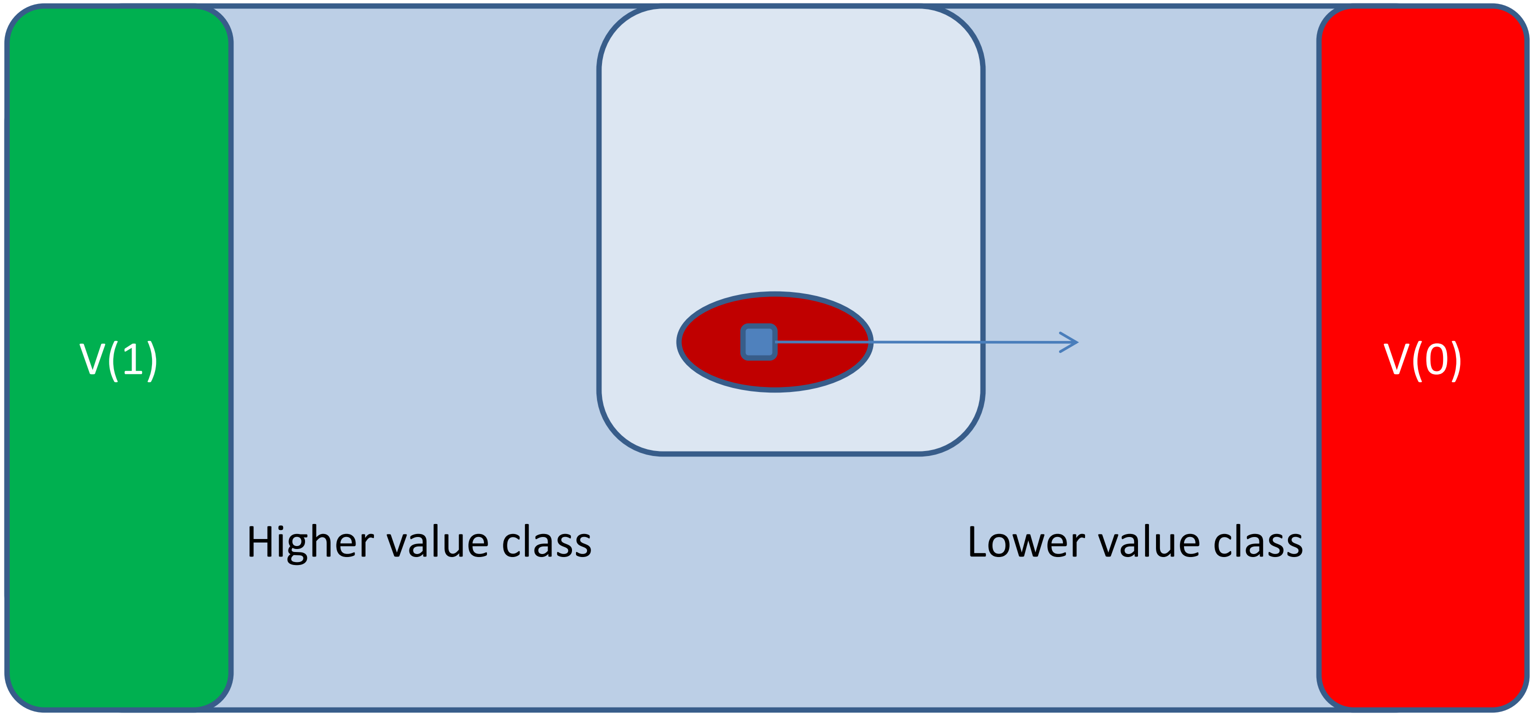
Sub-game Qualitative Optimality

- Claim: Player 1 wins almost-surely.
- Proof: Suppose not.
  - Then player 2 wins with positive probability somewhere.
  - Player 2 wins almost-surely somewhere.
  - Player 1 if stays in the value class loses with probability 1 or else jumps to a lower value class.
  - Contradiction.

Value Class: Boundary Probabilistic States



Value Class: Boundary Probabilistic States



Value Class Property

- In value classes if we assume boundary probabilistic vertices winning for player 1 then player 1 wins almost surely.
- Conditional almost-sure winning strategies.
- Stitching lemma: Compose them to get a optimal strategy.

Stitching Lemma

- Proof idea:
  - If the game stays in some value class player 1 wins with probability 1.
  - Else it leaves the value class through the boundary probabilistic vertex or goes to a higher value class.
  - Invoke sub-martingale Theorem or use results from MDPs.

Quantitative Analysis

- Pure memoryless optimal strategies exist.
- Complexity bound
  - NP $\cap$ coNP.
- Algorithms: Strategy improvement algorithms, uses qualitative algorithms and local optimality.

Stochastic Games Summary

	Reachability	**Liveness**	**Parity**
Qualitative	O(n m)	O(n m)	NP ∩ coNP Linear reduction to non-stochastic parity
Quantitative	NP ∩ coNP	NP ∩ coNP	NP ∩ coNP

Summary and Messages

- Markov chains
  - Qualitative: Linear time algorithm through closed recurrent states (bottom scc's).
  - Quantitative analysis: Linear equalities, Gaussian elimination.
- MDPs
  - Qualitative: Iterative algorithm.
  - Quantitative: Reduction to quantitative reachability using end-components.
  - Quantitative reachability: Linear programming.
- Stochastic games
  - Qualitative: Reduction to non-stochastic games.
  - Quantitative: Qualitative and local optimality.

Extensions

- Perfect-information turn-based finite state stochastic games
  - Infinite state games: pushdown games, timed games.
  - Concurrent games: simultaneous interaction.
  - Imperfect-information games.

CONCURRENT GAMES

Games on Graphs

Games on graphs:

1. Turn-based:
   - Chess.
   - Tic-tac-toe.

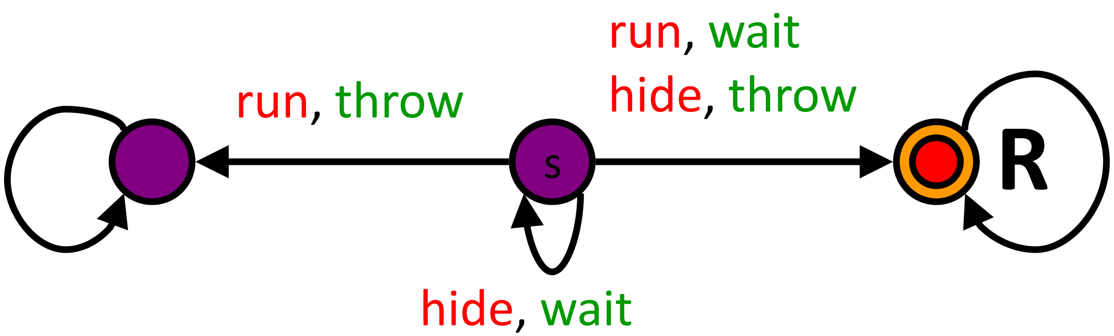
2. Concurrent:
   - Penalty Shoot-out.
   - Rock-paper-scissor.

Concurrent Game Graphs

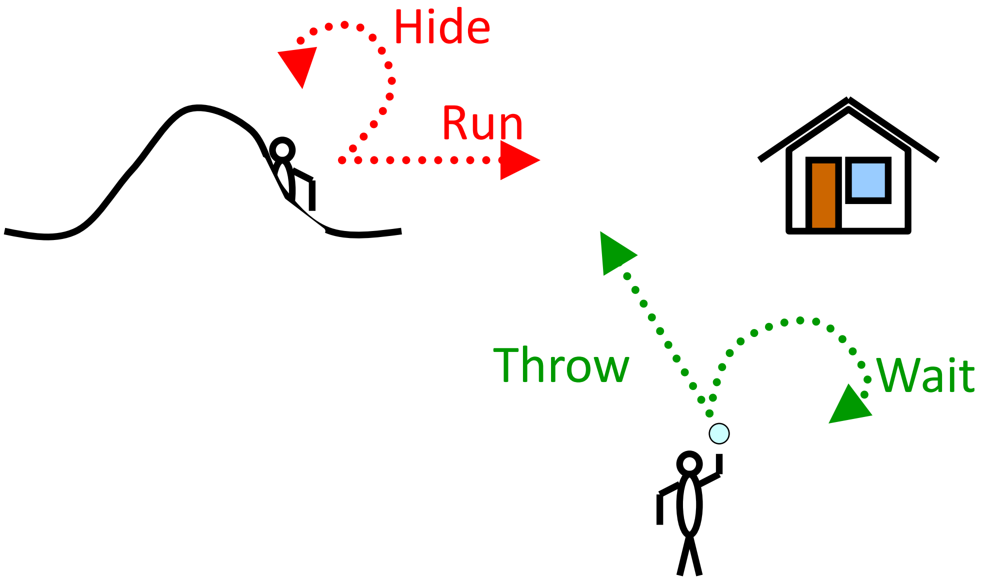
A concurrent game graph is a tuple G =(S,M,Γ_1,Γ_2,δ)

- S is a finite set of states.
- M is a finite set of moves or actions.
- Γ_i: S $\rightarrow$ $2^M \setminus \emptyset$ is an action assignment function that assigns the non-empty set Γ_i(s) of actions to player i at s, where i $\in$ {1,2}.
- δ: S $\times$ M $\times$ M $\rightarrow$ Dist(S), is a probabilistic transition function that given a state and actions of both players gives a probability distribution of the next state.

An Example (Deterministic Transition)

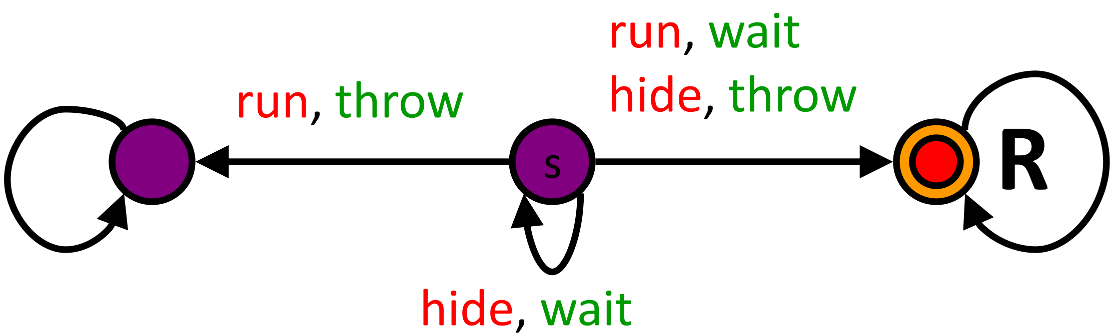


[Everett 57]



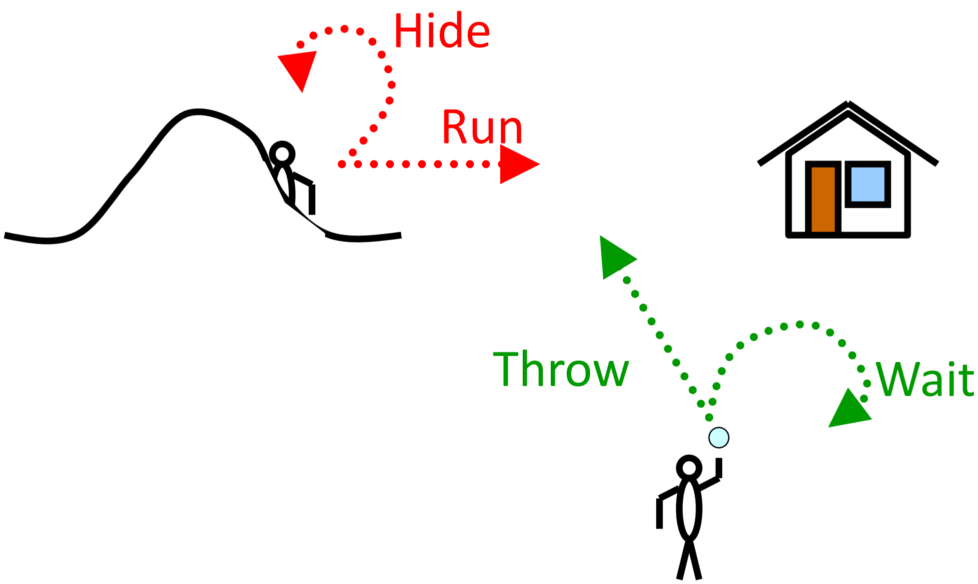
Concurrent reachability games

[Everett 57]

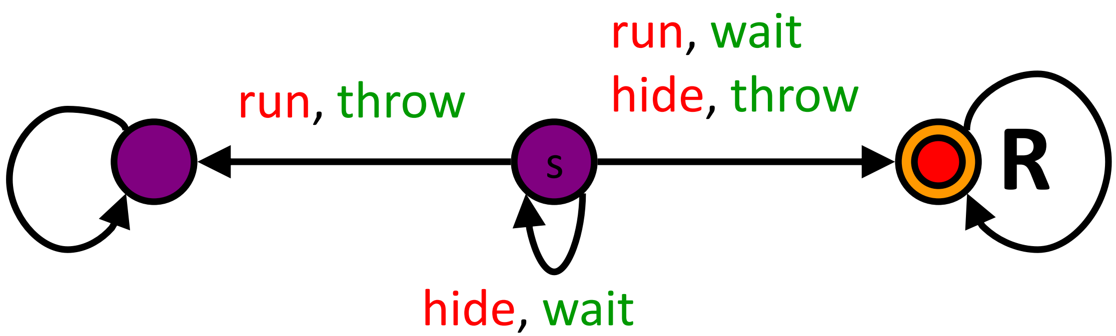


Move	Probability	
run	q	
hide	1-q	(q>0)

Win at s with probability
1-q, for all q > 0.

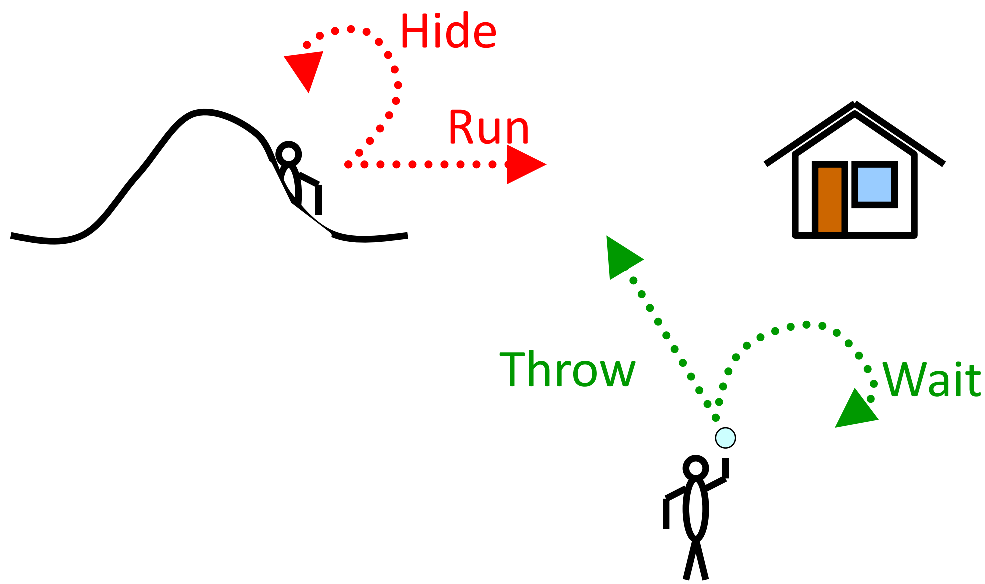


Concurrent reachability games



[Everett 57]

Move	Probability	
run	q	
hide	1-q	(q>0)



Win at s with probability
1-q, for all q > 0.

Player 1 cannot achieve v(s) = 1, only v(s) = 1-q for all q > 0.

Concurrent Games

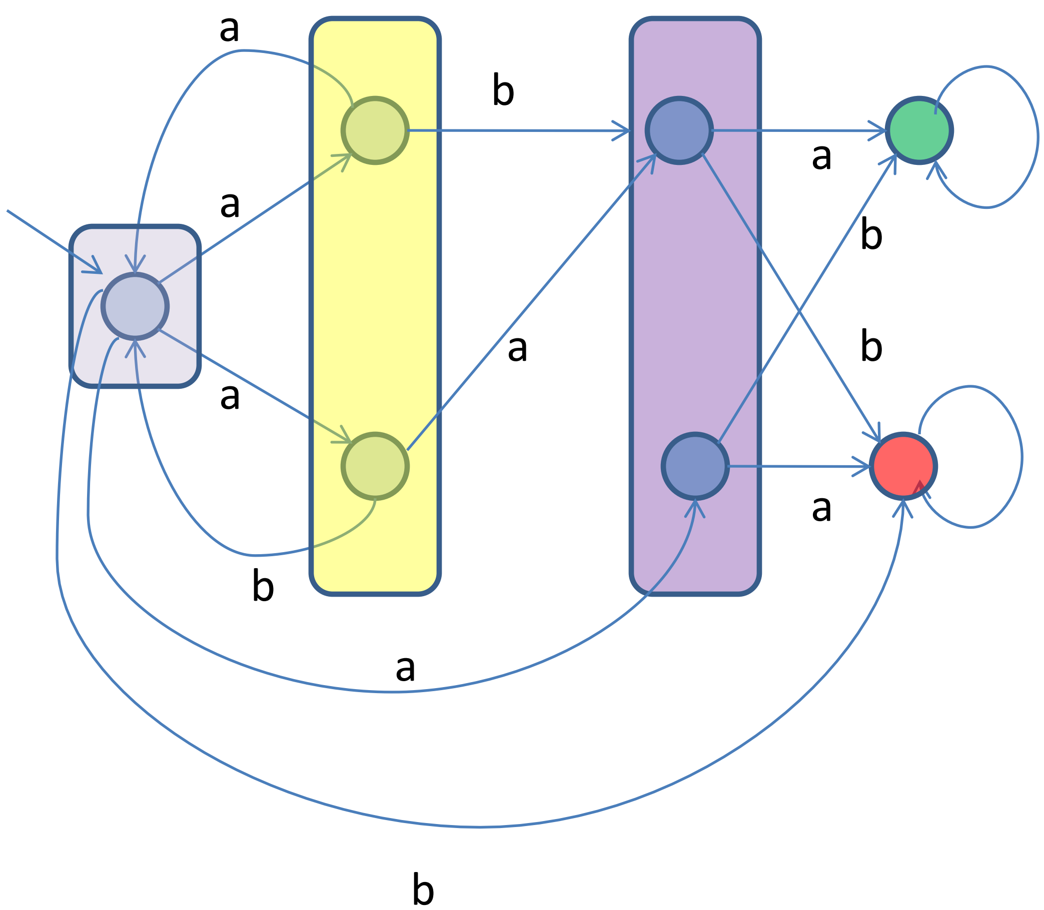
- Strategies
  - Require randomization.
  - May not be optimal.
  - Only ϵ-optimal, for $\epsilon>0$.
  - For liveness requires infinite memory.

- Values can be irrational for concurrent deterministic reachability games.

- Qualitative and quantitative analysis still decidable
  - Qualitative analysis is NP $\cap$ coNP.
  - Quantitative analysis is PSPACE.

PARTIAL-INFORMATION GAMES

Partial-information Games



In starting play a.
In yellow play a and b at random.
In purple:

- if last was yellow then a
- if last was starting, then b.

Requires both randomization and memory

Partial-information Games

- Strategies
  - Require randomization.
  - May not be optimal.
  - Only ϵ-optimal, for ϵ>0.
  - For liveness requires infinite memory.
  - More complicated than concurrent games.

- Quantitative analysis
  - Undecidable.

- Qualitative analysis
  - Reachability, Liveness: EXPTIME-complete.
  - Parity: Undecidable.

Conclusion

- Perfect-information stochastic games
  - Applications: verification and synthesis of stochastic reactive systems.
  - Markov chains, MDPs and stochastic games with parity objectives.

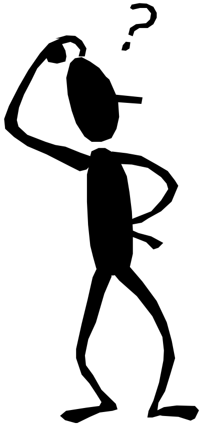
- Glimpses of the world of games beyond.

References

- Applications and connections:
  - Church 62, Pnueli-Rosner 89, Ramadge-Wonham 87, Courcoubetis-Yannakakis 95, Thomas Handbook 97, and many more.
- Markov chains:
  - Book of Kemeny
- Markov Decision Processes:
  - Book of Filar Vrieze.
  - Probabilistic verification: Courcoubetis-Yannakakis.
  - PhD Thesis of deAlfaro.
- Stochastic games
  - Condon 92, 93.
  - PhD Thesis of Chatterjee

The end

Thank you !



Questions ?