Some Complexity Results for Subclasses of Stochastic Games

Krishnendu Chatterjee

I|S|T AUSTRIA
Institute of Science and Technology

Workshop on Stochastic Games,
Singapore, Nov 30, 2015

Stochastic Games

- This talk glimpse of two types of results:
  - Computational complexity.
  - Strategy complexity.
  - For stochastic games as well as many different subclasses.

Stochastic Game Graphs

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

- S is a finite set of states.

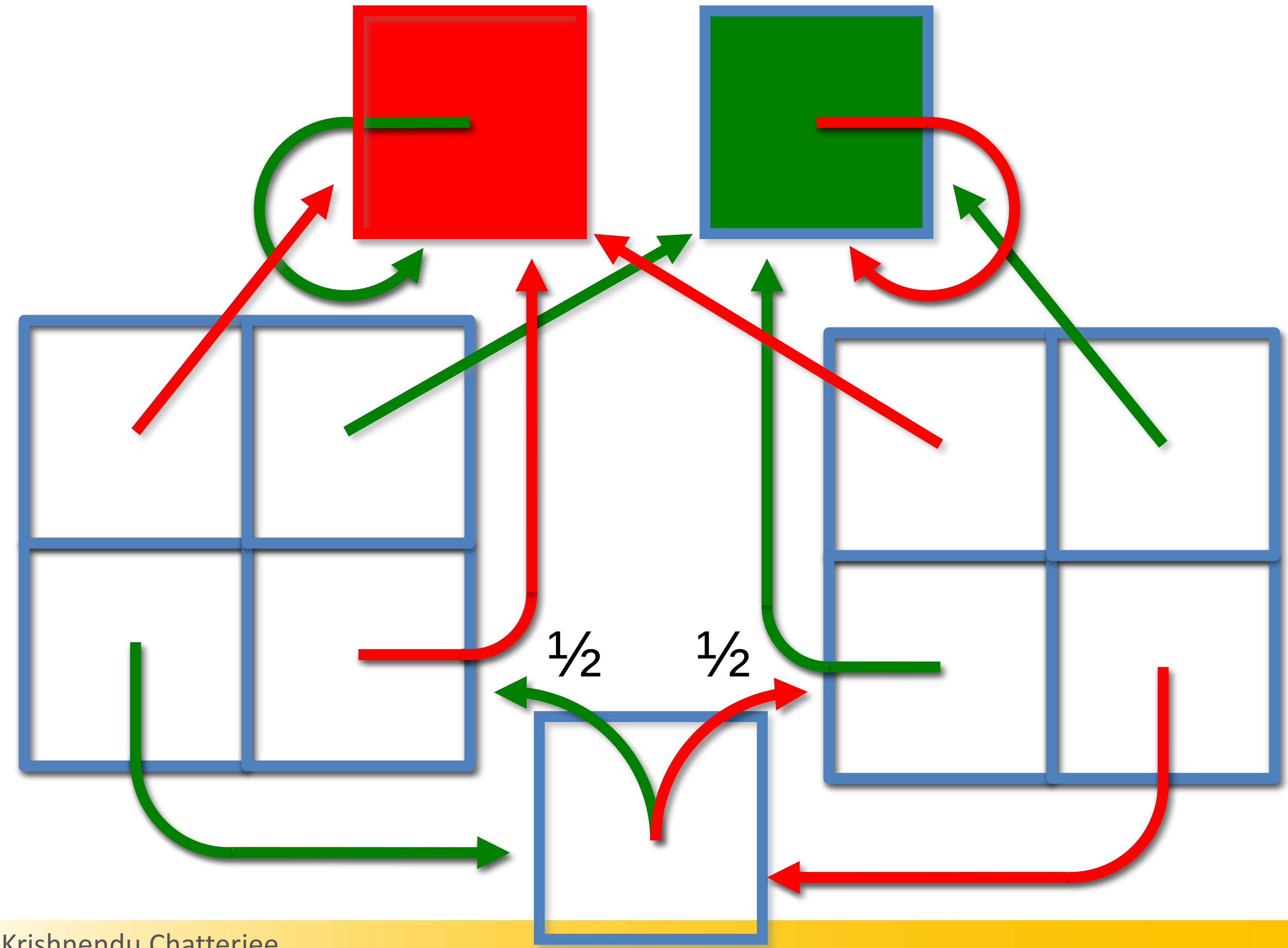
- M is a finite set of moves or actions.

- 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}.

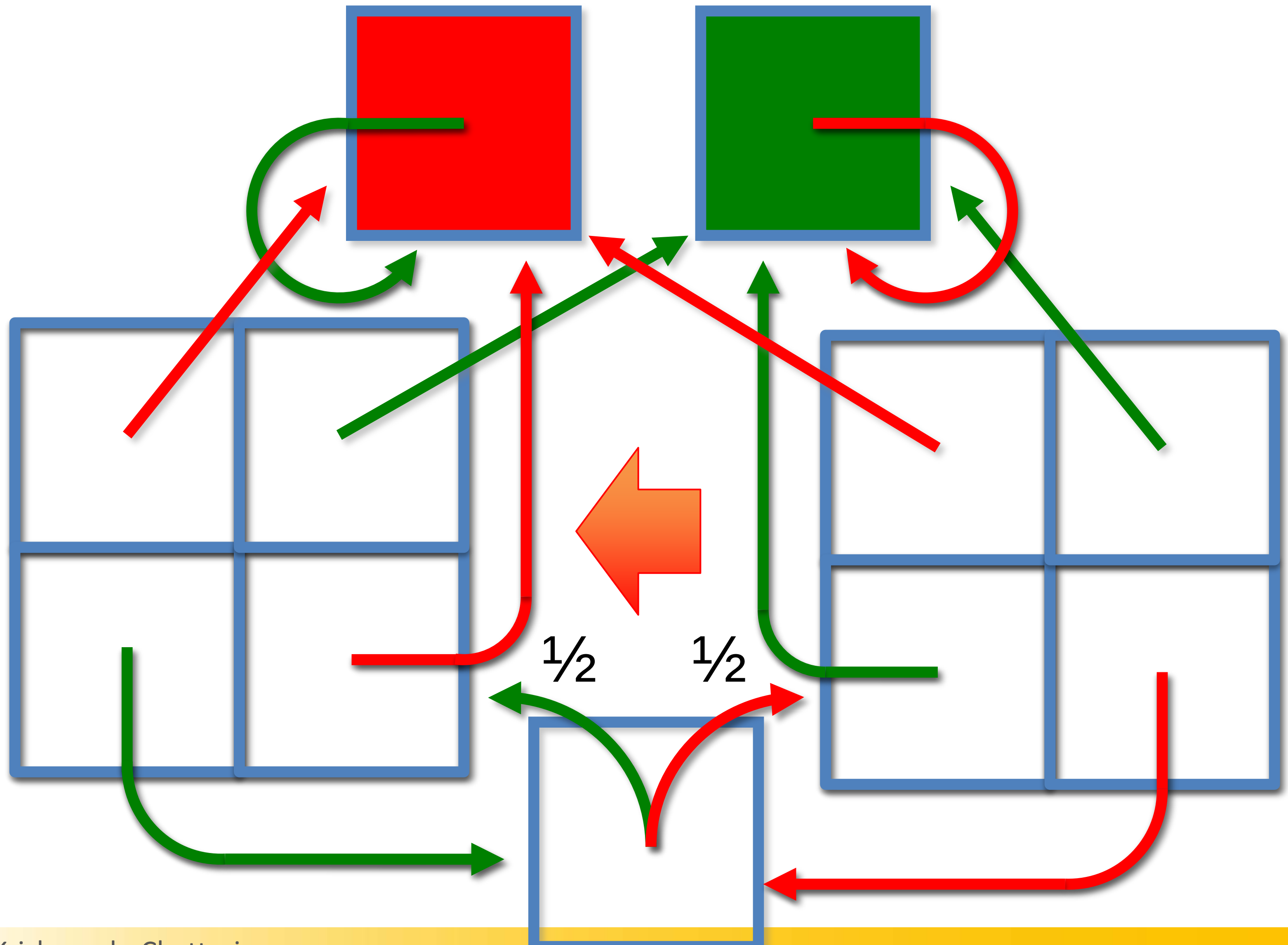
- δ: S $\times$ M $\times$ M $\rightarrow$ D(S), is a *stochastic* transition function that given a state and actions of both players gives a distribution over the next state.

- For deterministic games, the transition function is *deterministic*.

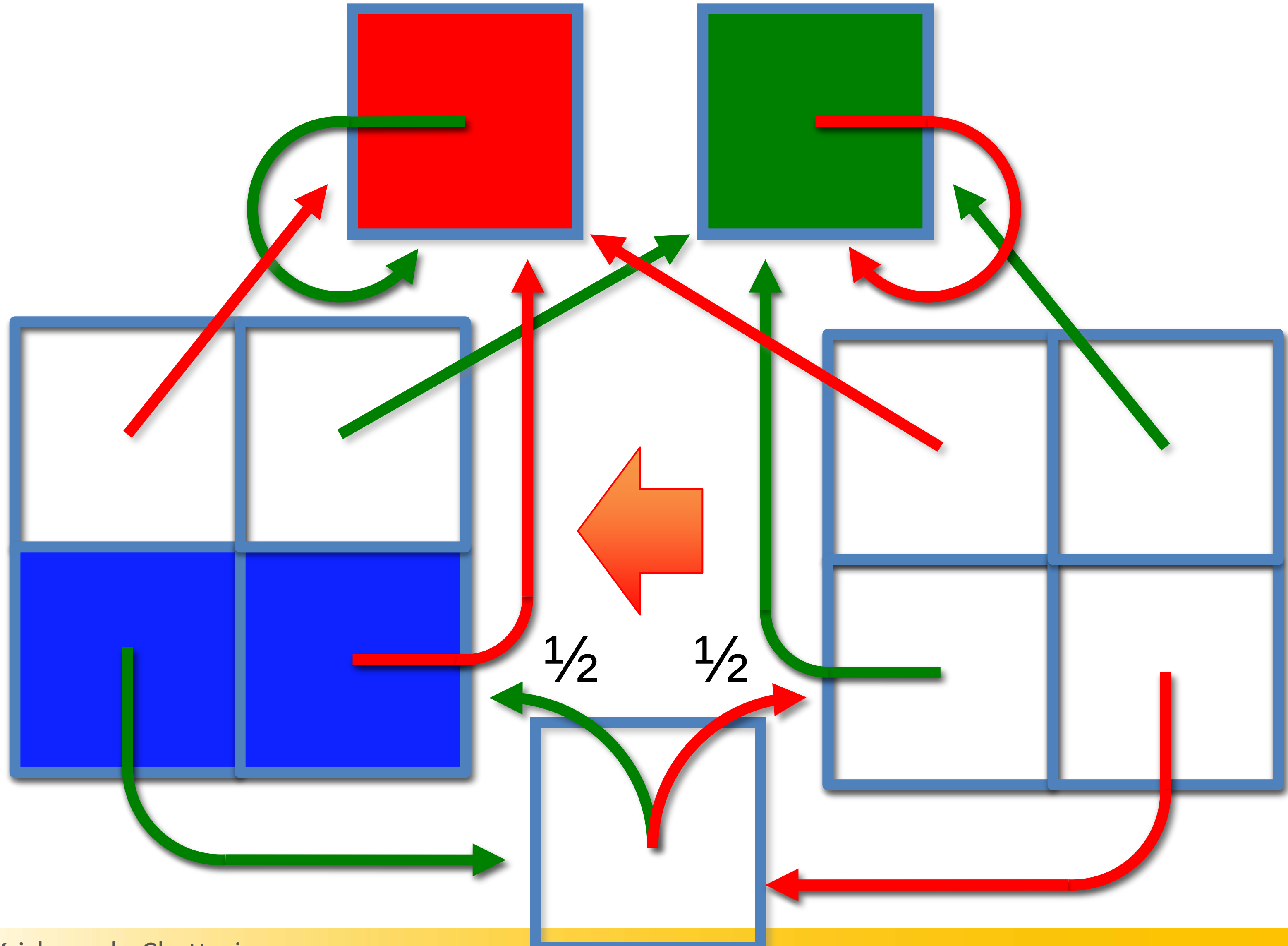
Example game



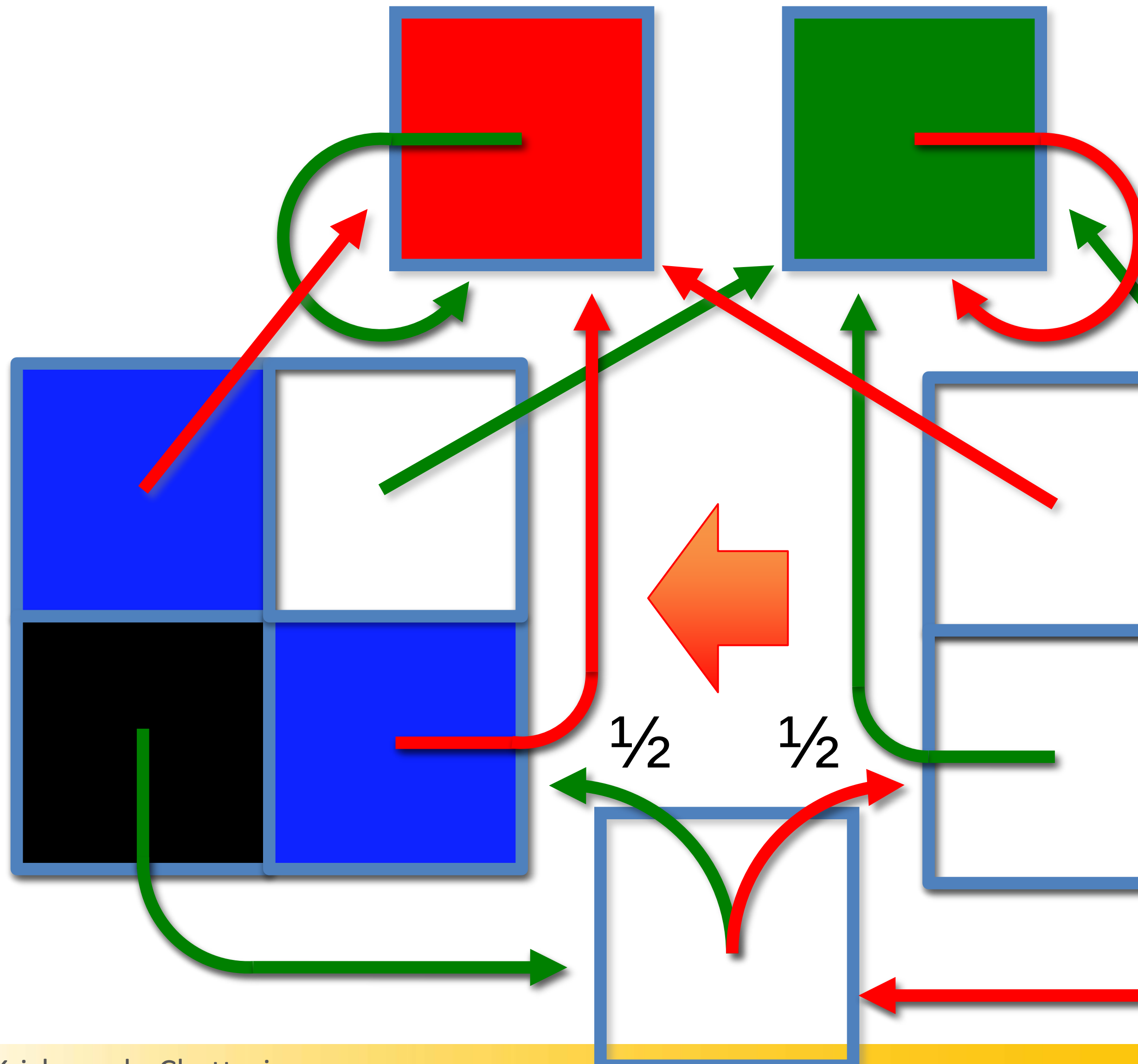
Example game

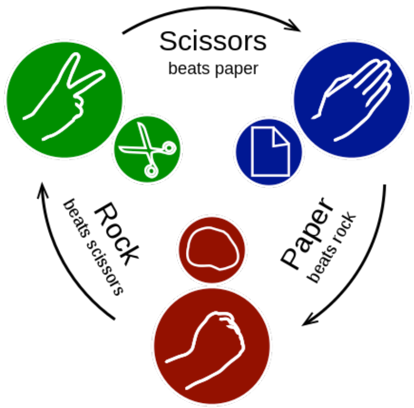


Example game

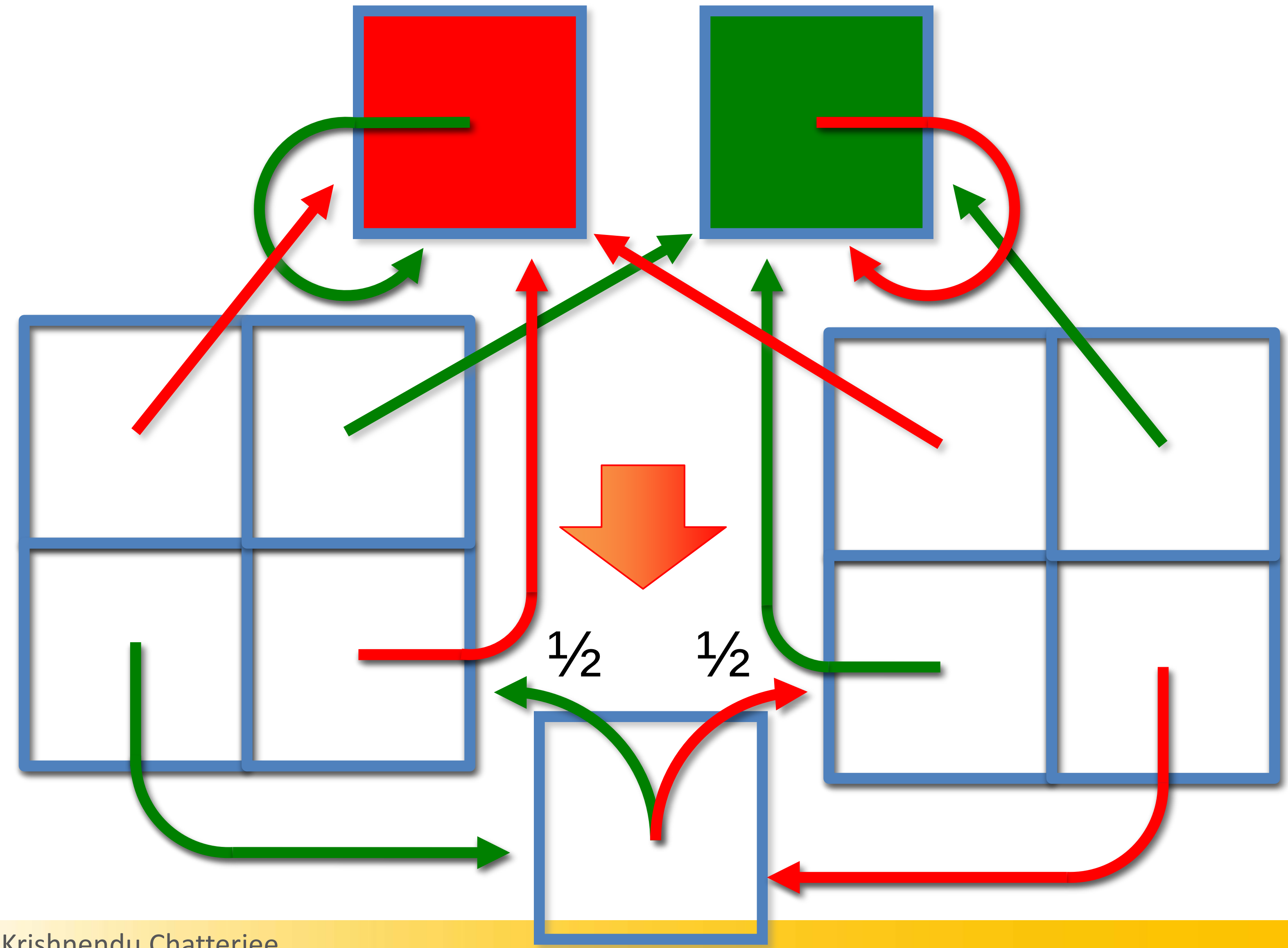


Example game

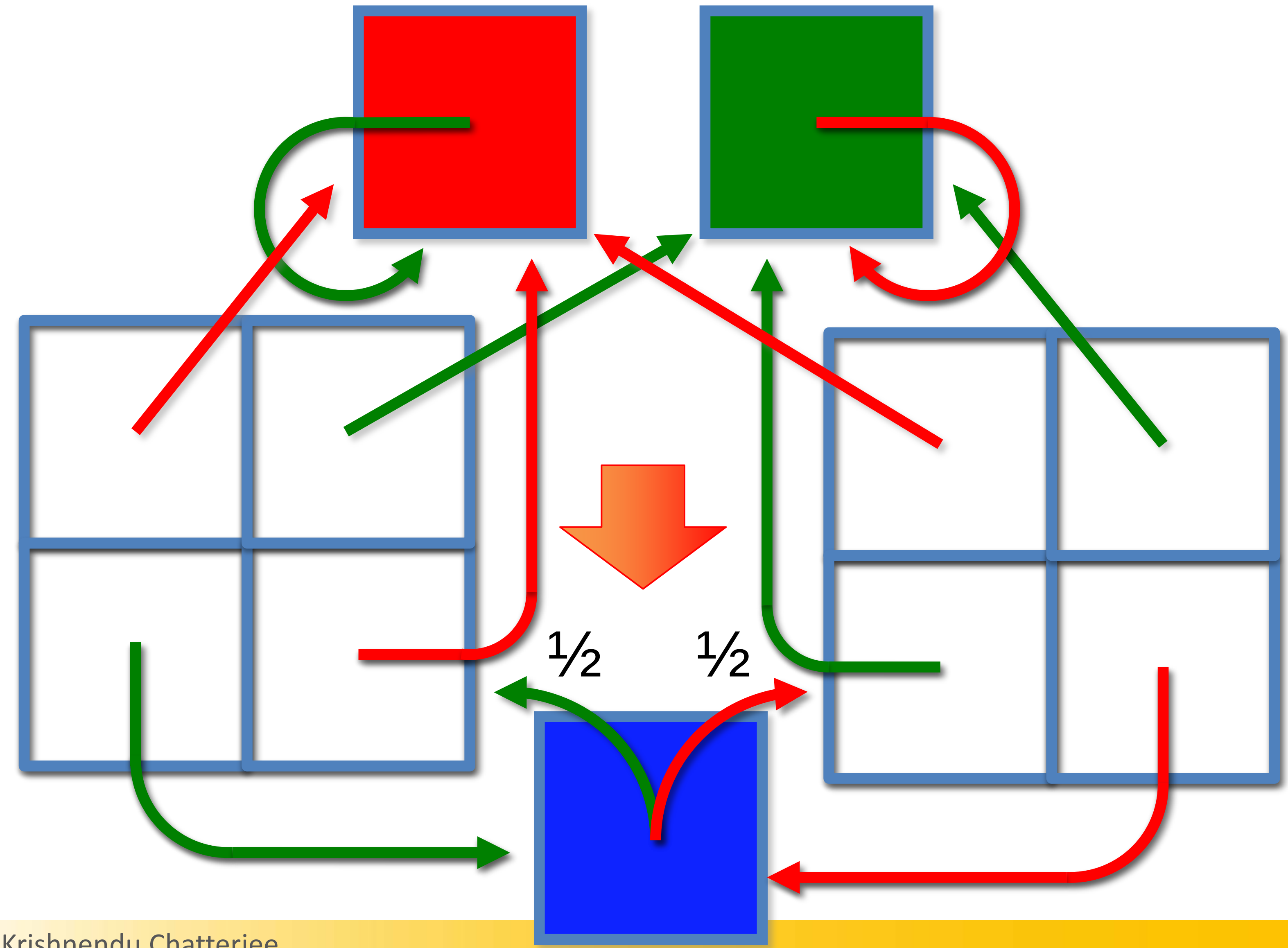




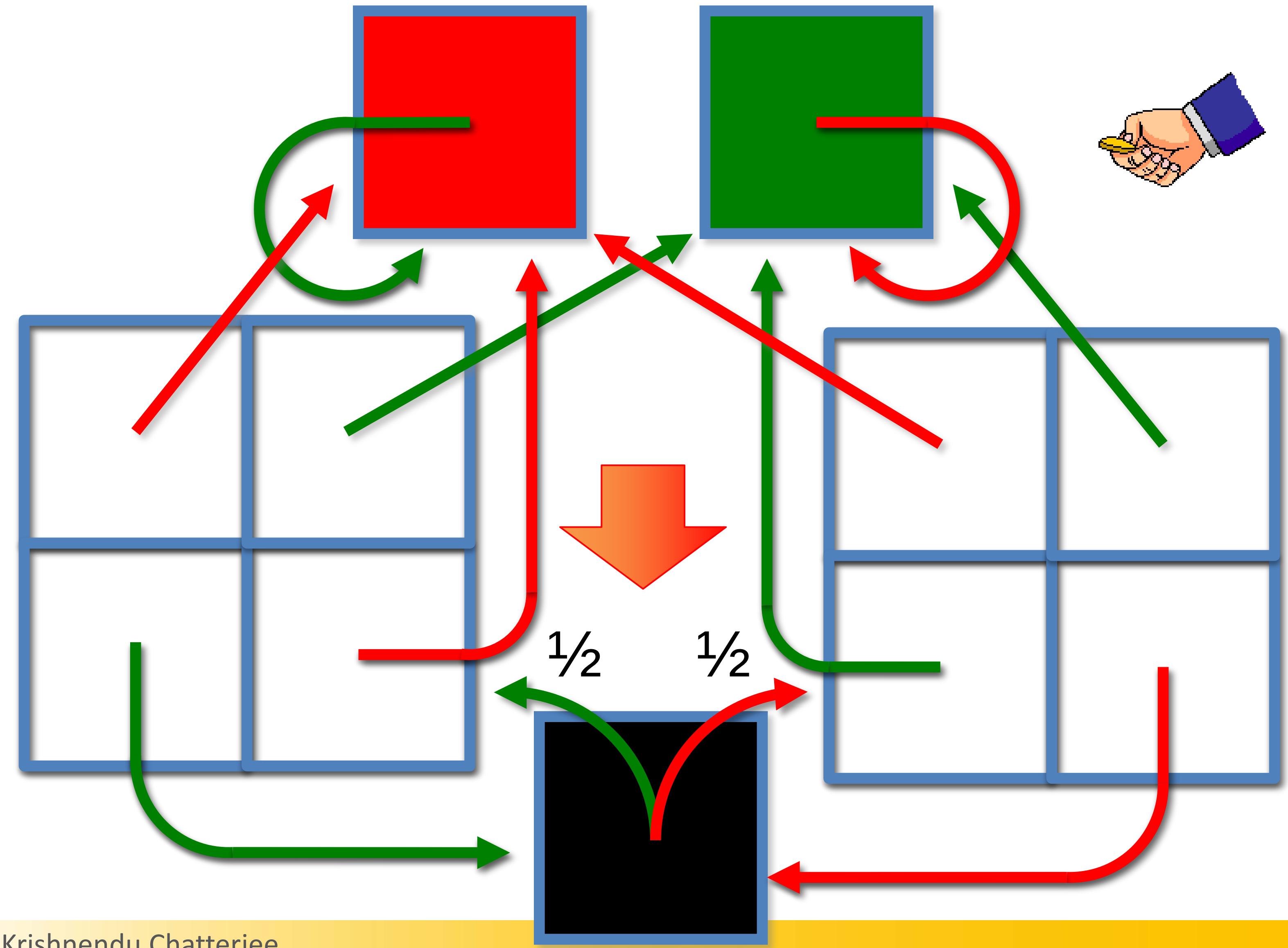
Example game



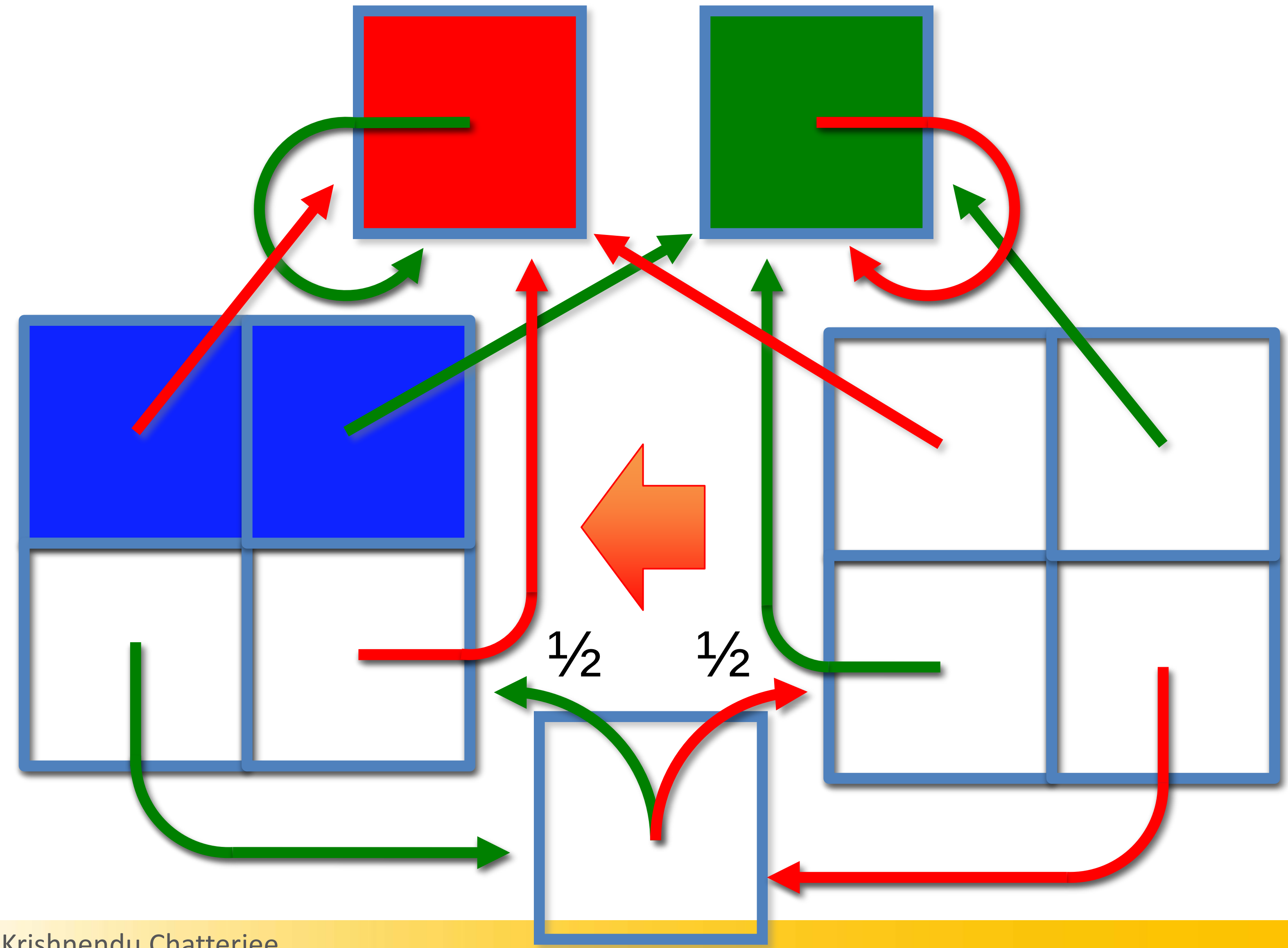
Example game



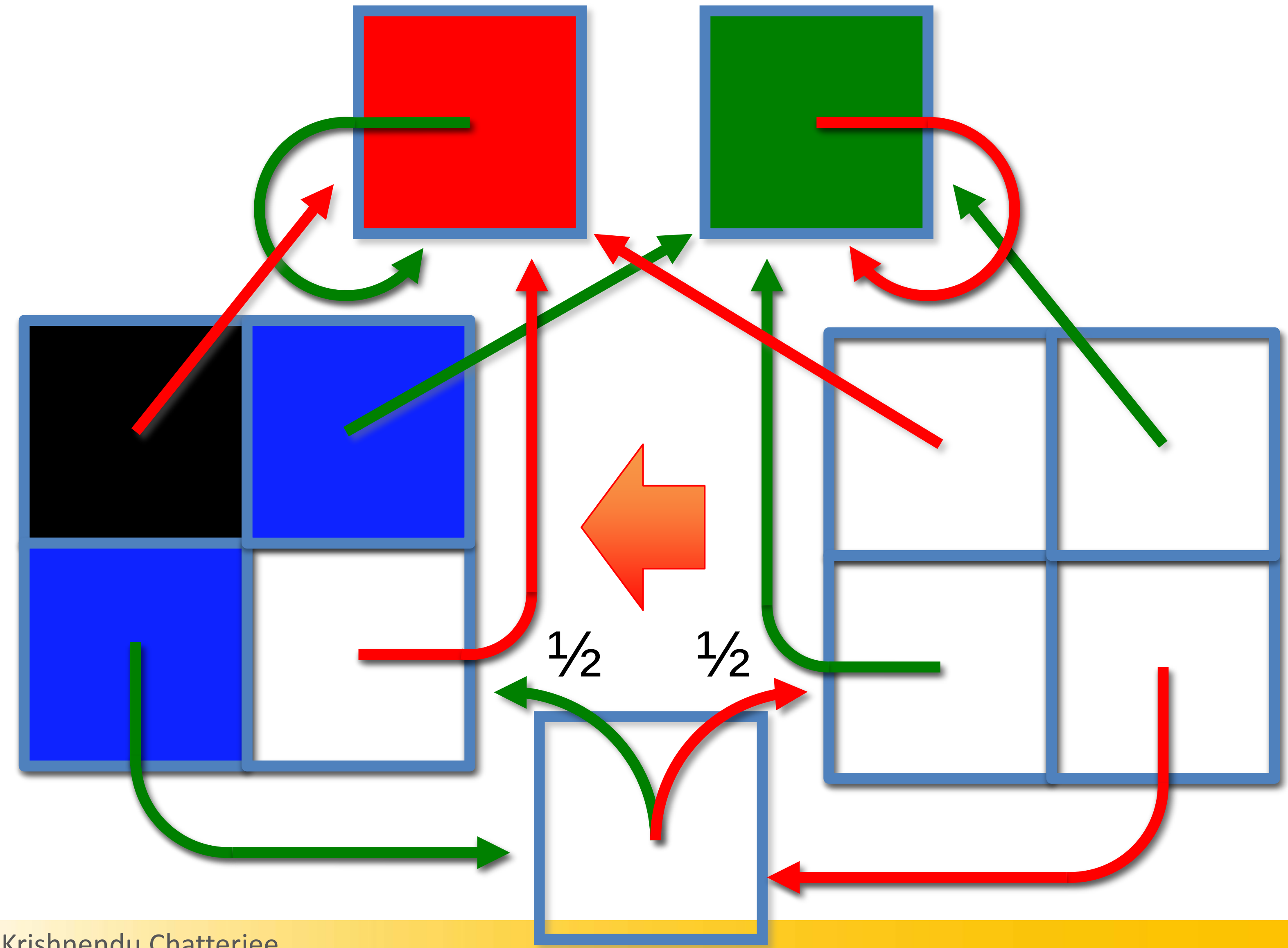
Example game



Example game



Example game



Strategies

- Recipes to play the game.
  - σ: $(S \cdot M \cdot M)^* \cdot S \rightarrow D(M)$
- Complexity of strategies:
  - Memory.
  - Randomization.
- Stationary strategies (no memory):
  - σ: $S \rightarrow D(M)$

Mean-payoff Objective

- Every transition is assigned a rational reward in the interval [0,1], by a reward function r.

- Mean-payoff objective: The payoff for a play (infinite path) is the long-run average of the rewards of the path.
  - LimSupAvg.
  - LimInfAvg.

Existence of Value

- Fundamental result on existence of values [MN81]
  - $\sup_{\sigma} \inf_{\pi} E_{s, \sigma \pi}$ [LimInfAvg] = $\inf_{\pi} \sup_{\sigma} E_{s, \sigma \pi}$ [LimSupAvg]
  - Order of strategies can be exchanged.
  - The value of the game v(s).
- Value problem: The basic computational problem is to decide whether v(s) $\geq \lambda$.

Survey of Results

- Computational complexity of the value problem.
- Strategy complexity: Strategies for witness of the value problem.
- General stochastic games and various subclasses.

General Problem Result

- Decision problem:
  - First result:
    - Exponential time: $2^{poly(m,n)}$, where m is number of actions, and n is number of states [CMH08].
  - Second result:
    - Doubly exponential: m^{2^n} [HKLMT 11].
    - For constant number of states is polynomial. Nice generalization of zero-sum matrix games.
  - Strategy complexity: very complicated even for simple games like Big-match.

Towards Subclasses

Concurrent games, Mean-payoff obj

Towards Subclasses

Concurrent games, Mean-payoff obj

Structural restr.

Turn-based stochastic

Ergodic

Turn-based deterministic

Towards Subclasses

Concurrent games, Mean-payoff obj

Structural restr.

Objective restr.

Turn-based stochastic

Ergodic

Reach obj.

Turn-based deterministic

Reachability and Safety Games

- Reachability/safety games:
  - A set T of terminal or absorbing states with reward 1, all other states have reward 0.
  - Hence the reachability player wishes to reach T, and safety player wishes to avoid T.
- Most basic objectives in computer science
  - Reactive safety critical systems.
- Positive recursive games
  - Reachability player.
  - Safety player is the opponent.

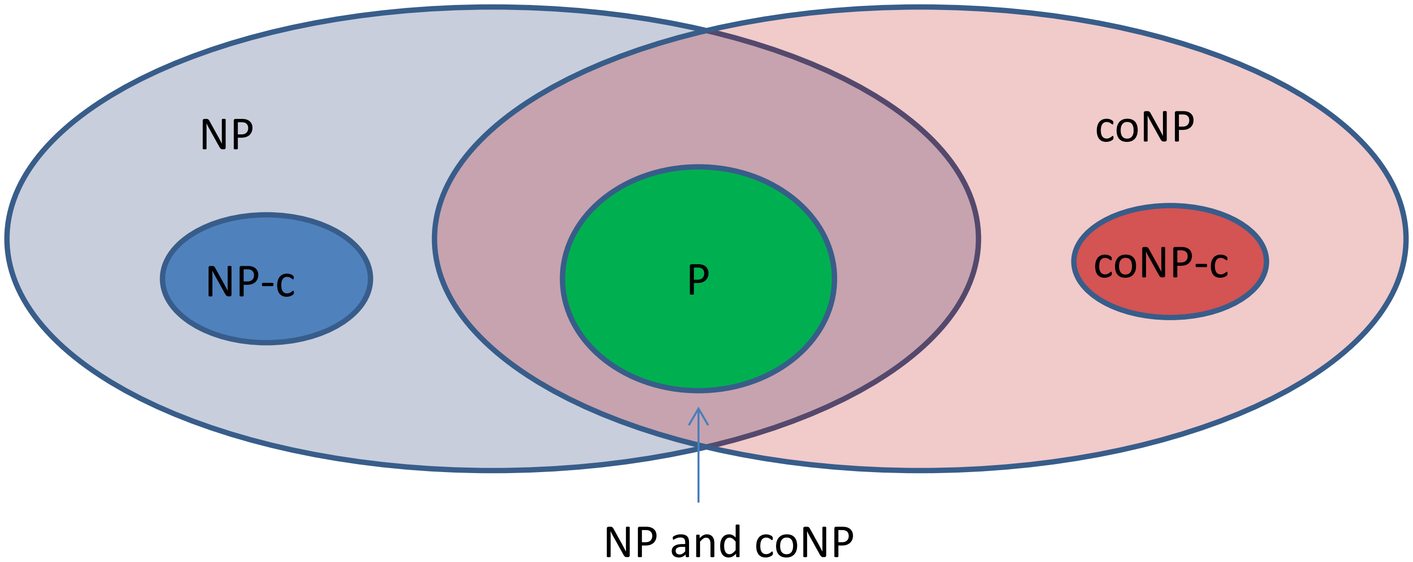
Computational Classes

- Polynomial time (*P*): Efficient
  - Linear, Quadratic.

- Non-deterministic polynomial time (*NP*):
  - Given a witness of polynomial length it can be checked in polynomial time.

- *coNP* some sense complement of *NP*
  - Given a counter-witness (to show some answer is no) of polynomial length it can be checked in polynomial time.

Computational Classes



TURN-BASED (STOCH. & DET.) GAMES

Turn-based Games: Computational Complexity

1. **Turn-based deterministic:**
   a) Reach: Linear time.
   b) Mean-payoff [EM79,ZP95,Karp79]:
      I. O(n m W);
      II. *NP and coNP*; not known to be *P*.

2. **Turn-based stochastic:**
   a) Reach:
      I. *NP and coNP*, not known to be *P*.
      II. At least as hard as 1b [Con92].

   b) Mean-payoff:
      I. Equivalent to 2a [AM09].

Turn-based Games: Computational Complexity

1. Turn-based deterministic:
   a) Reach: Linear time.
   b) Mean-payoff [EM79,ZP95,Karp79]:
      I. O(n m W);
      II. *NP and coNP*; not known to be *P*.

2. Turn-based stochastic:
   a) Reach:
      I. *NP and coNP*, not known to be *P*.
      II. At least as hard as 1b [Con92].

   b) Mean-payoff:
      I. Equivalent to 2a [AM09].

Turn-based Stochastic Games

- Strategy complexity [LL69]:
  - Positional (deterministic and stationary).
- The *NP and coNP* bound:
  - Polynomial witness:
    - Positional strategy.
    - An action for every state.
  - Polynomial time verification:
    - Given a positional strategy is fixed we obtain an MDP.
    - Values in MDPs can be computed in polynomial time by linear programming [FV97].

Some Hardness Results

- Hardness results:
  - TBD Mean-payoff Value Problem.
  - TBS Reach Value Problem.
  - SQUARE-ROOT-SUM problem:
    - Given positive integers a_1, a_2, … , a_n, and b, decide if the sum of square roots of a_i is at least b.
    - This problem is not even known to be in NP.

ERGODIC GAMES

Ergodic Games

- For all strategies all states appear infinitely often with probability 1.

- Stationary optimal strategies exist [HK66].
  - However, not positional, randomization is need.

- Strategy complexity of stationary strategies
  - How complex is to *represent* the probability distribution of a stationary strategy.

Stationary Strategy Representation

- Distribution in every state.
- Representation of distributions
  - Exponential numbers have polynomial-size representation due to binary representation.
  - Doubly exponential numbers cannot be explicitly represented in polynomial size.
  - Distributions that can be expressed with exponential numbers have polynomial representation.

Stationary Strategies Complexity

- Complexity measure:
  - Patience: Inverse of minimum non-zero probability [Eve57].
  - Roundedness: The number r such that all probabilities multiple of 1/r.
  - Pat $\leq$ Rou.
  - Significance:
    - Exponential roundedness implies polynomial witness.
    - Doubly exponential patience implies explicit representation requires exponential space (not polynomial witness in explicit representation).

Ergodic Games Results [CI 14]

- Reachability is not relevant.

- Strategy complexity:
  - For ϵ-optimal strategies, for ϵ>0, we show exponential patience is necessary (lower bound) and exponential roundedness is sufficient (upper bound).

  - Lower bound based on a family of games.

  - Upper bound based on a coupling argument.

Ergodic Games Results [CI 14]

- Computational complexity:
  - Value problem (precise decision question): is SQUARE-ROOT-SUM hard.
  - Value problem (precise or approximate): TBS Value problem hard.
  - Approximation problem is in NP.

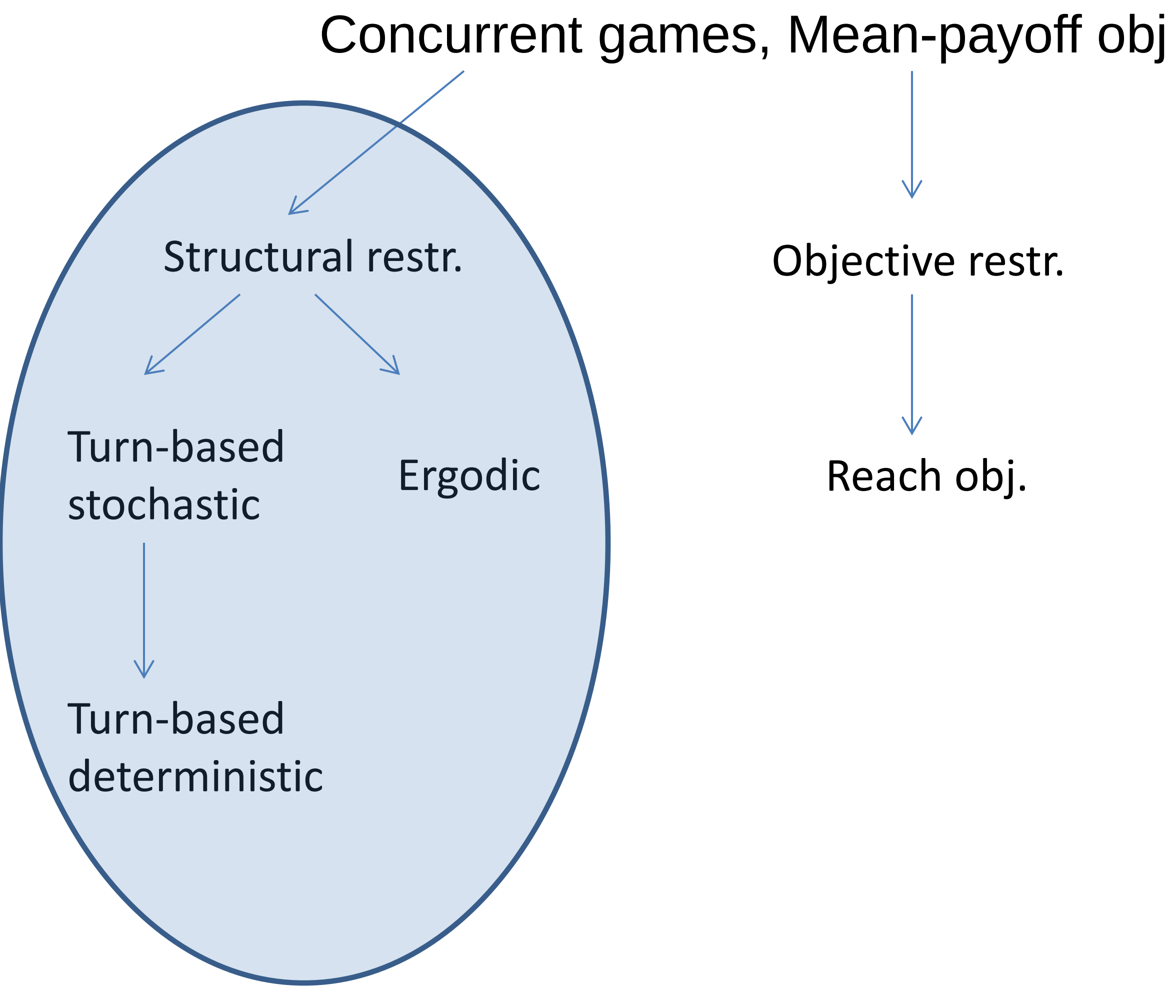
Ergodic Games Results [CI 14]

- Strategy complexity of optimal strategies:
  - We don't know a precise answer.
  - We have the following result: Exponential patience for optimal strategies would imply SQUARE-ROOT-SUM problem in *P*.
  - Hence proving exponential patience will be a major breakthrough. Proving super-exponential lower bound would separate optimal and ϵ-optimal strategies.

Summary of Results

	TB Det	TB Stoch Value	Conc. Ergodic Value
Reach	Linear	*NP and coNP* Open ques: in *P*	----
Mean-payoff	*NP and coNP* Open ques: in *P*	*NP and coNP* Open ques: in *P*	*NP and coNP* (approx) Hardness (approx) SQRT-SUM-hard (exact)

Towards Subclasses



CONCURRENT REACH/SAFE GAMES

Reachability and Safety Games

- Reachability/safety games:
  - A set T of terminal or absorbing states with reward 1, all other states have reward 0.
  - Hence the reachability player wishes to reach T, and safety player wishes to avoid T.
- Positive stochastic games
  - Reachability player.
  - Safety player is the opponent.

Reachability and Safety Games

- Computational complexity:
- Value problem
  - Exponential time: [dAM01].
  - SQUARE-ROOT-SUM –hard: [EY06].
  - Approximation problem: NP^{NP} [FM13].

Reachability and Safety Games

- Strategy complexity:

- Reachability player [Eve57]:
  - Optimal strategies need not exist, but ϵ-optimal for all ϵ>0.
  - ϵ-optimal strategies, for ϵ>0, are stationary.

- Safety player [Par71]:
  - Optimal stationary strategies exist.
  - Locally optimal strategies are optimal.

Reachability and Safety Games

- Strategy complexity:

- Reachability player results.
  - Doubly-exponential patience is necessary and doubly-exponential roundedness is sufficient [HKM09].

Reachability and Safety Games

- Strategy complexity: Reachability/safety player comparison (based on number of value classes). New results [CHI15].

# Value classes	Reachability	Safety
1	Linear	One
2	Double-exponential	One
3	Double-exponential	
Constant	Double-exponential	
General	Double-exponential	

Reachability and Safety Games

- Strategy complexity: Reachability/safety player comparison (based on number of value classes). New results [CHI15].

# Value classes	Reachability	Safety
1	Linear	One
2	Double-exponential	One
3	Double-exponential	**Exponential**
Constant	Double-exponential	**Exponential**
General	Double-exponential	

Reachability and Safety Games

- Strategy complexity: Reachability/safety player comparison (based on number of value classes). New results [CHI15].

# Value classes	Reachability	Safety
1	Linear	One
2	Double-exponential	One
3	Double-exponential	**Exponential**
Constant	Double-exponential	**Exponential**
General	Double-exponential	**Double-exponential**

Reachability and Safety Games

- Strategy complexity: Reachability/safety player comparison (based on number of value classes). New results [CHI15].

# Value classes	Reachability	Safety
1	Linear	One
2	Double-exponential	One
3	Double-exponential	**Exponential**
Constant	Double-exponential	**Exponential**
General	Double-exponential	**Double-exponential**

Surprising result

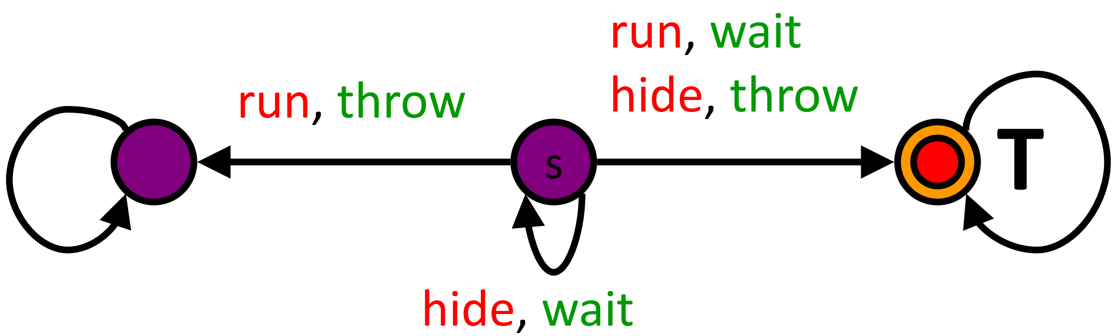
Surprising Results

- 3-state lower bound
  - Two terminal state and one state.
  - Local optimally implies optimality. So basically play strategies of matrix games.
  - In matrix games, only logarithmic patience is necessary.
  - For safety games, in matrix, there is a variable, which depends on the value. This causes an increase from logarithmic to exponential.

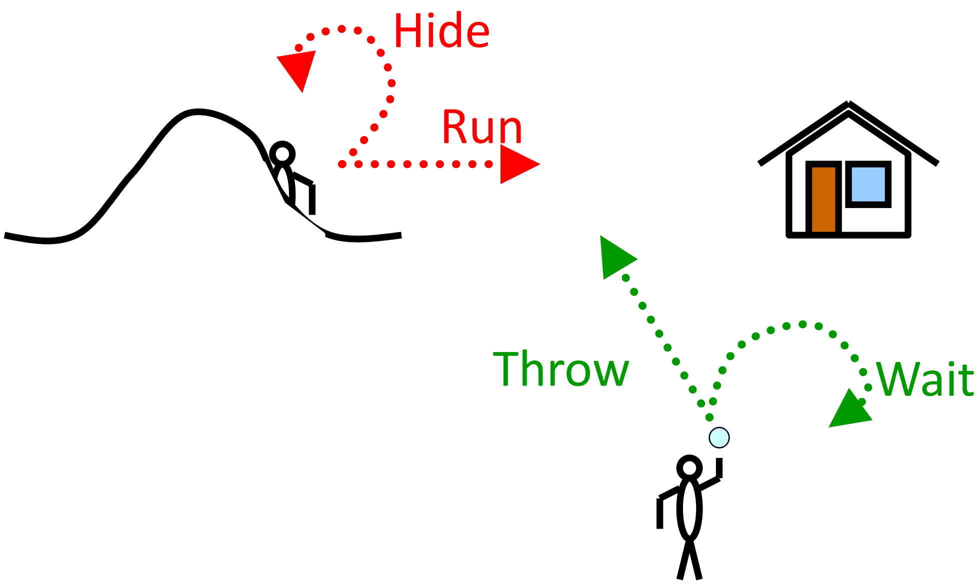
The Doubly Exponential LB

- Lower bound for safety is surprising:
  - Two other games which share properties with safety.
  - Discounted games: Local optimality implies optimality and there exponential roundedness suffices.
  - Ergodic games: optimal stationary strategies exist, and again exponential roundedness suffices.
- First explain the lower bound for reachability.
- Then the lower bound for safety.

An Example: Snow-ball Game [dAHK98]

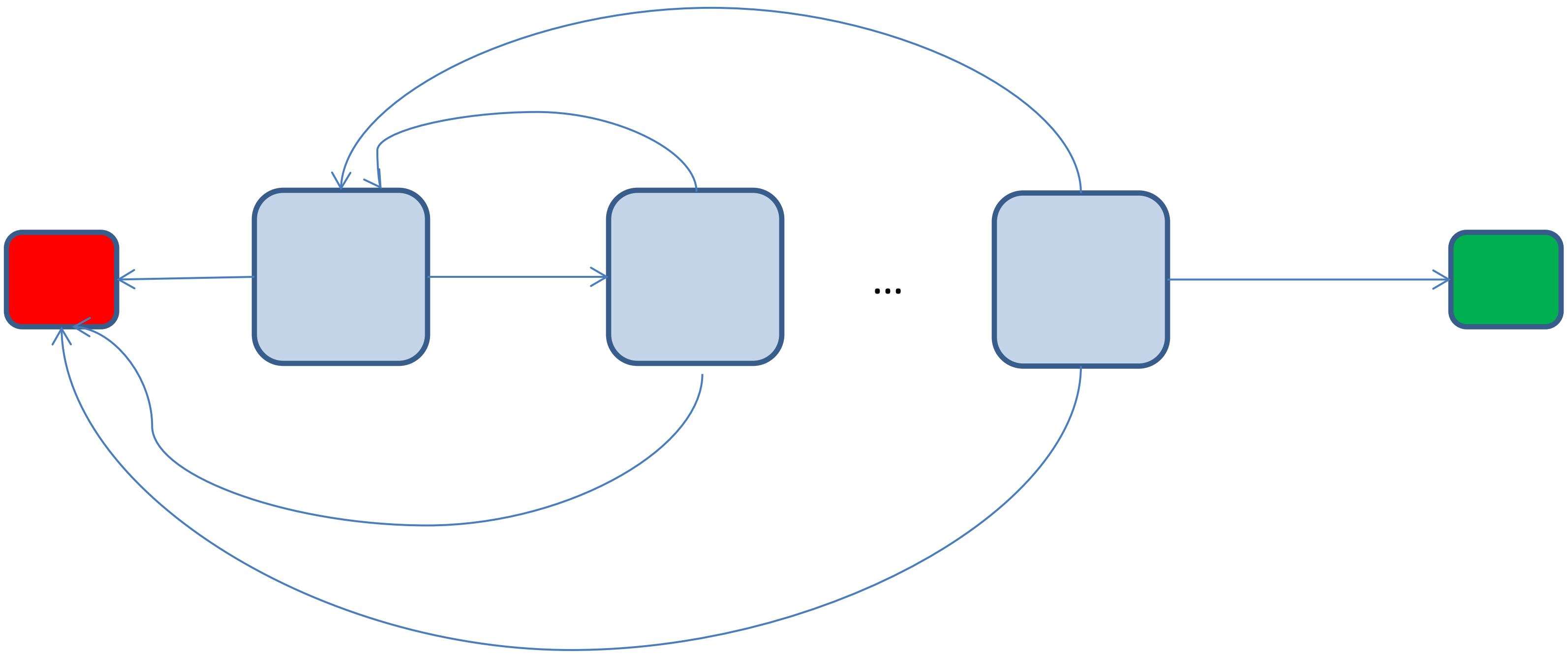


[Eve 57]



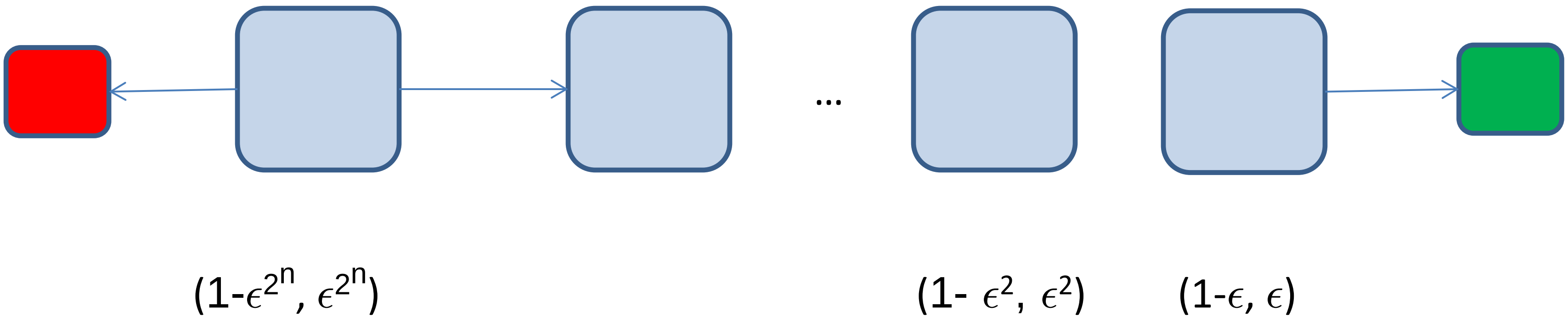
Play hide 1-ϵ,
Run ϵ

Snow-ball-in Stages: Purgatory [HKM09]



Success event: Move forward one step.
Mistake event: Loose the game.
Stay event: Back to the start state.
To remove cluttering will omit the arrows in next slides.

Snow-ball-in Stages: Purgatory [HKM09]

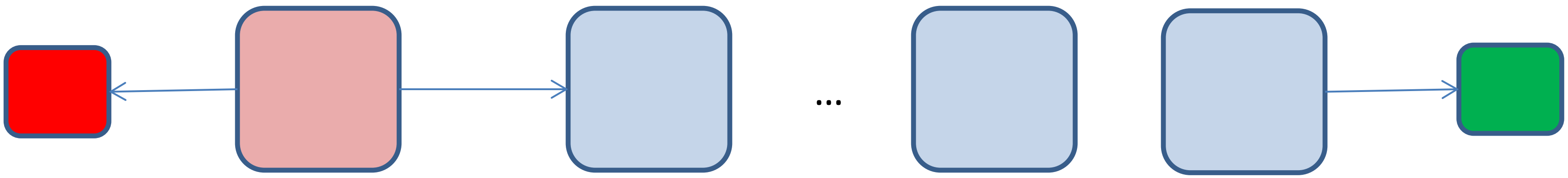


Reachability player: Doubly exponential patience is necessary.

In this game, the safety player has positional optimal strategies.

We will call this game Pur(n): n stages.

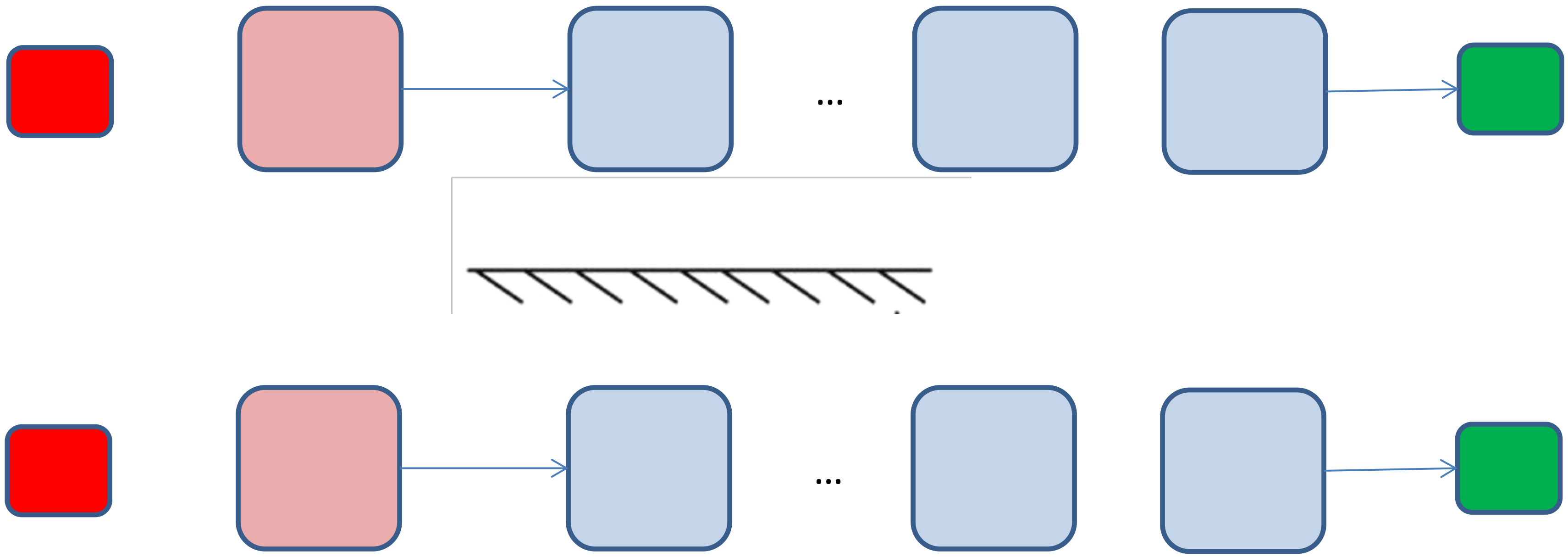
Towards the Safety Game Counter Example



1. Consider Pur(n+1).

2. Simplify the start state by making it deterministically go to the next state. SimPur(n).

Towards the Safety Game Counter Example



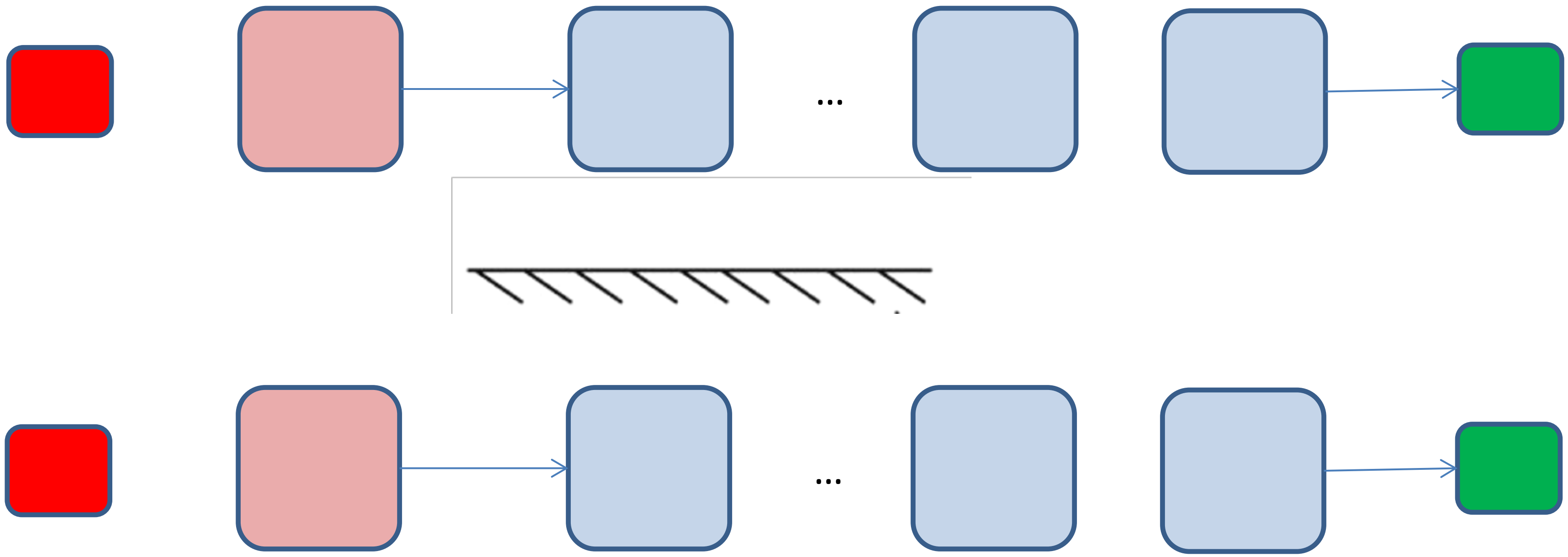
2. SimPur(n).

3. Take its mirror image. Exchange role of players. MirSimPur(n)

Towards Safety Game Counter Example

- SimPur(n): Safety player has positional strategies.
- MirSimPur(n): Safety player has positional strategies.

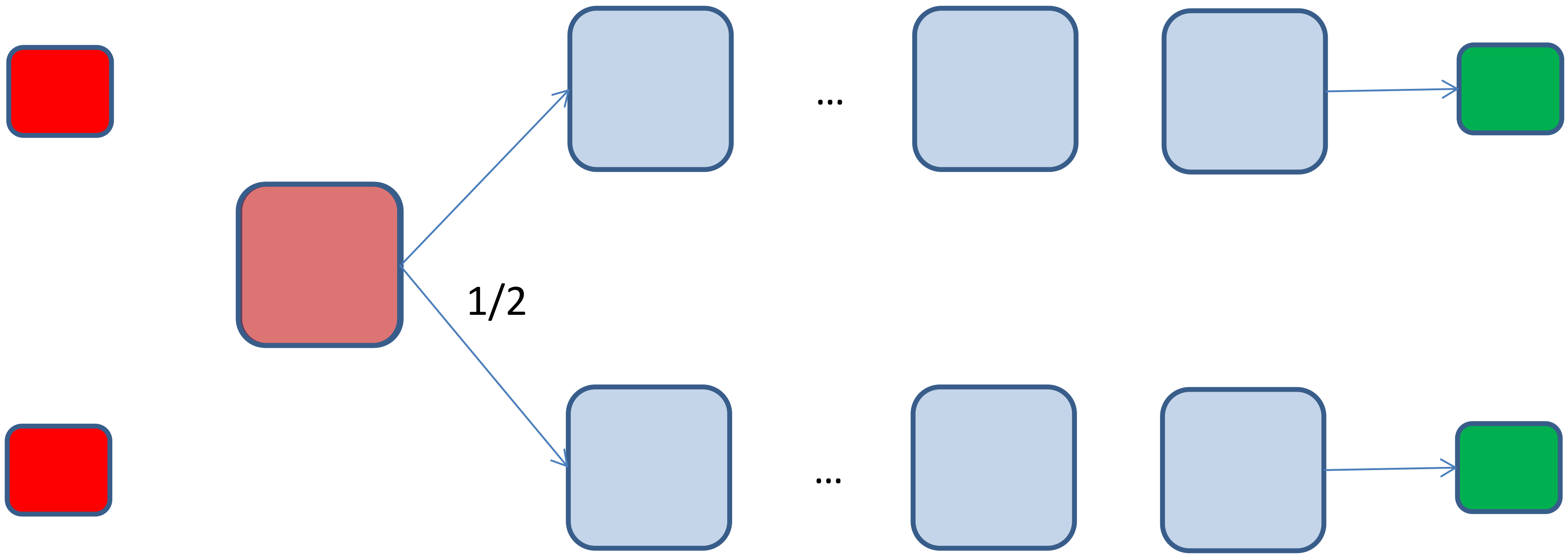
Towards the Safety Game Counter Example



2. SimPur(n).

3. MirSimPur(n)

Towards the Safety Game Counter Example



2. SimPur (n).

3. MirSimPur(n)

4. Merge start states. PurDuel(n)

Towards Safety Game Counter Example

- PurDuel(n): Safety player requires doubly exponential patience.
- Merging two games where positional suffices we get a game where doubly exponential patience is necessary.

Summary: Concurrent Reachability and Safety Games

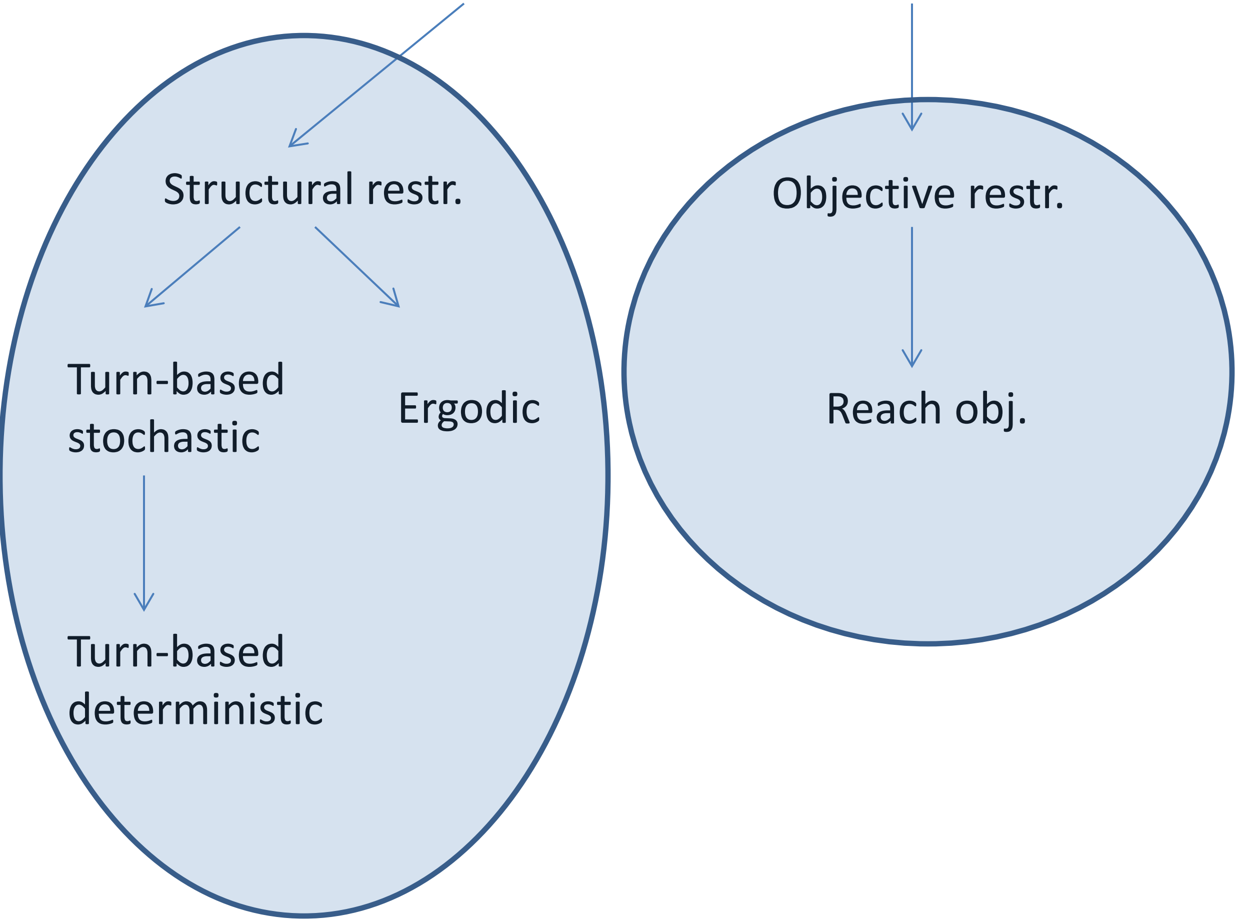
- Computational complexity: Value problem
  - Exponential time (polynomial space): [dAM01].
  - SQUARE-ROOT-SUM –hard: [EY06].
  - Approximation problem: NP^{NP} [FM 13].

- Strategy Complexity:

# Value classes	Reachability	Safety
1	Linear	One
2	Double-exponential	One
3	Double-exponential	**Exponential**
Constant	Double-exponential	**Exponential**
General	Double-exponential	**Double-exponential**

Towards Subclasses

Concurrent games, Mean-payoff obj



CONCLUSION AND OPEN PROB

Conclusion

- Strategy and computational complexity of the value problem for stochastic games.

- Two restrictions:
  - Structural: Turn-based, ergodic.
  - Objective: Reachability.

- Other restrictions:
  - Value-1 problem.
  - Special classes of strategies.

- Survey of results:
  - Some polynomial time, some open questions.

Major Open Questions

- Value problem for TBD Mean-payoff in *P*.
- Value problem for TBS reach games in *P*.

Collaborators

- Kristoffer Arnsfelt Hansen
- Thomas A. Henzinger
- Rasmus Ibsen-Jensen
- Rupak Majumdar

References

- [MN81]

J. Mertens and A. Neyman. Stochastic games. IJGT, 10:53–66, 1981.

- [CMH08]

K. Chatterjee, R. Majumdar, and T. A. Henzinger. Stochastic limit-average games are in EXPTIME. IJGT, 37(2):219–234, 2008.

- [HKLMT11]

K. A. Hansen, M. Koucky, N. Lauritzen, P. B. Miltersen, and E. P. Tsigaridas. Exact algorithms for solving stochastic games: extended abstract. In STOC, pages 205–214, 2011.

- [EM79]

A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. IJGT, 8(2):109–113, 1979.

- [ZP96]

U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. Theoretical Computer Science, 158:343–359, 1996.

- [Con 92]

A. Condon. The complexity of stochastic games. I&C, 96(2):203–224, 1992.

- [AM09]

D. Andersson and P. B. Miltersen: The Complexity of Solving Stochastic Games on Graphs. ISAAC 2009: 112-121

References

- [HK66]

A. J. Hoffman and R. M. Karp. On nonterminating stochastic games. Management Science, 12(5):359–370, 1966.

- [Eve57]

H. Everett. Recursive games. In *CTG*, volume 39 of *AMS*, pages 47–78, 1957.

- [CI14]

K. Chatterjee and R. Ibsen-Jensen. The Complexity of Ergodic Mean-payoff Games. In *ICALP 2014*, pages 122–133, 2014.

- [MS07]

P. B. Miltersen and T. B. Sørensen. A near-optimal strategy for a heads-up no-limit texas hold'em poker tournament. In *AAMAS'07*, pages 191–197, 2007.

- [dAM01]

L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. In STOC'01, pages 675–683. ACM Press, 2001.

- [EY06]

K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. In *ICALP'06 (2)*, pages 324–335, 2006.

- [FM13]

S. K. S. Frederiksen and P. B. Miltersen. Approximating the value of a concurrent reachability game in the polynomial time hierarchy. In *ISAAC*, pages 457–467, 2013.

References

- [dAHK98]

L. de Alfaro, T. A. Henzinger, and O. Kupferman. Concurrent reachability games. FOCS, 1998.

- [Par 71]

T. Parthasarathy. Discounted and positive stochastic games. *Bull. Amer. Math. Soc*, 77:134–136, 1971.

- [HKM 09]

K. A. Hansen, M. Koucky, and P. B. Miltersen. Winning concurrent reachability games requires doubly-exponential patience. In LICS, pages 332–341, 2009.

- [CHI15]

K. Chatterjee, K. A. Hansen and R. Ibsen-Jensen: Strategy Complexity of Concurrent Stochastic Games with Safety and Reachability Objectives. CoRR abs/1506.02434 (2015).

- [CI 15a]

K. Chatterjee and R. Ibsen-Jensen: Qualitative analysis of concurrent mean-payoff games. I&C. 242: 2-24 (2015)

- [CI 15b]

K. Chatterjee and R. Ibsen-Jensen: The Value 1 Problem Under Finite-memory Strategies for Concurrent Mean-payoff Games. SODA 2015: 1018-1029.

- [HIK 15]

K. A. Hansen, R. Ibsen-Jensen and M. Koucky. Personal communication. For a copy contact Ibsen-Jensen.

QUESTIONS?