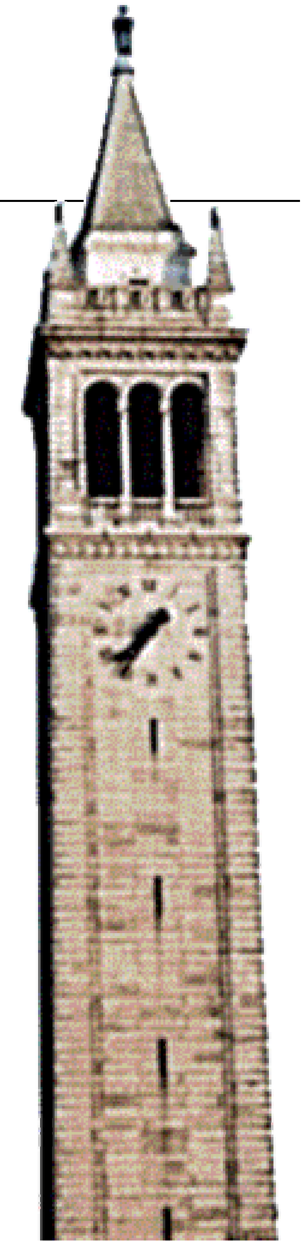
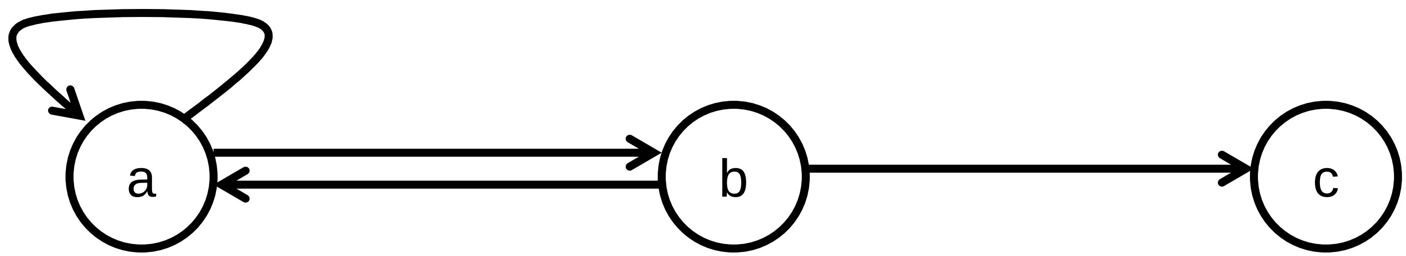
Discounting the Future in Systems Theory

Luca de Alfaro, UC Santa Cruz
Tom Henzinger, UC Berkeley
Rupak Majumdar, UC Los Angeles

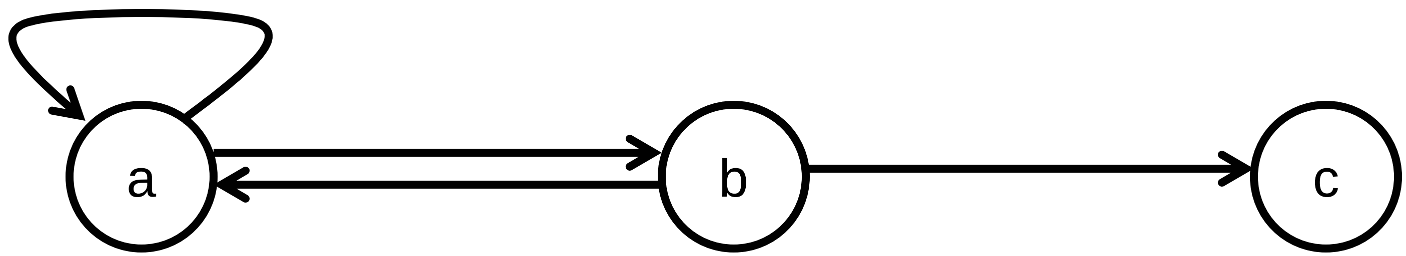
Chess Review
May 11, 2005
Berkeley, CA



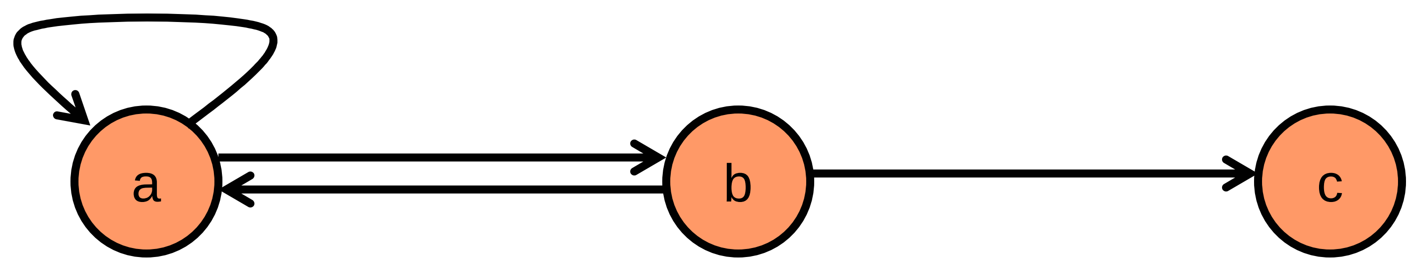
A Graph Model of a System



Property ◇c ("eventually c")

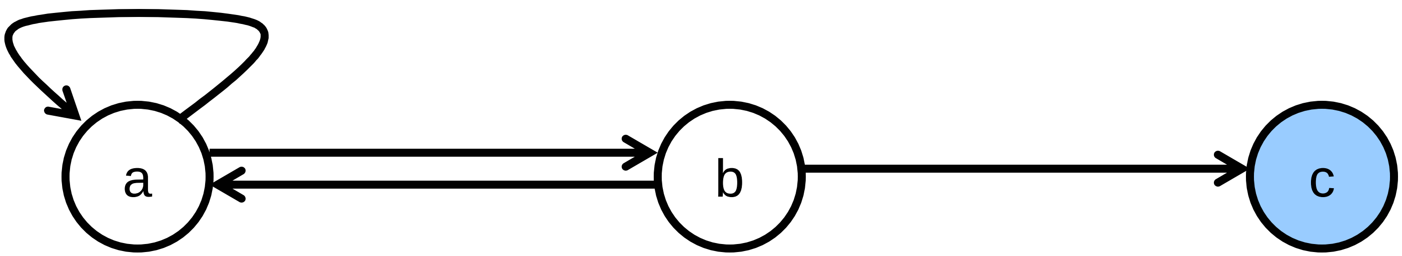


Property ◇c ("eventually c")



∃ ◇**c** … some trace has the property ◇c

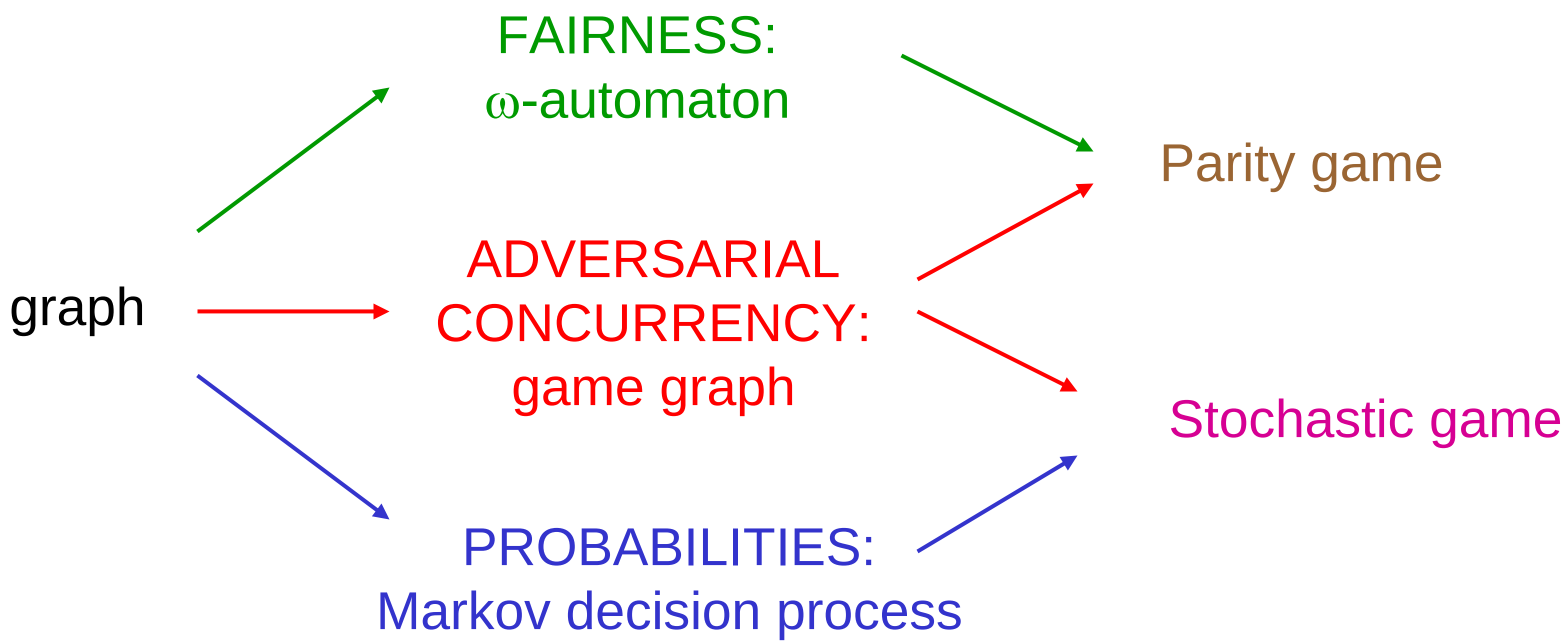
Property ◇c ("eventually c")



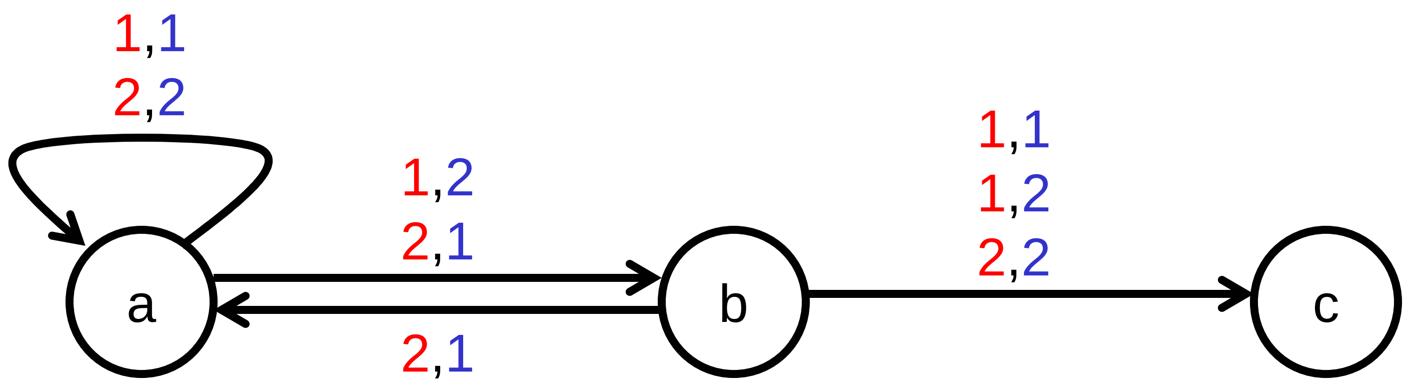
∃ ◇c ... some trace has the property ◇c

∀ ◇c ... all traces have the property ◇c

Richer Models



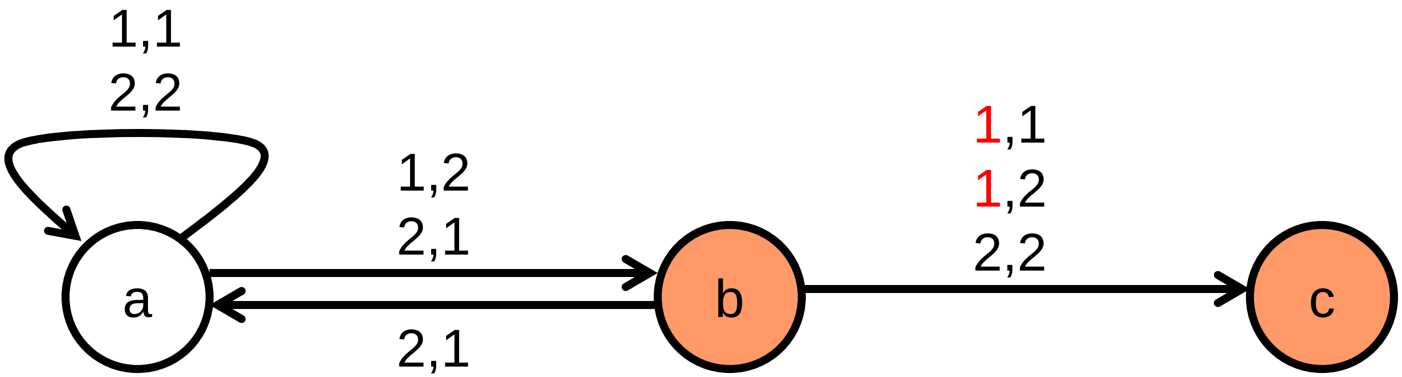
Concurrent Game



player "left"
player "right"

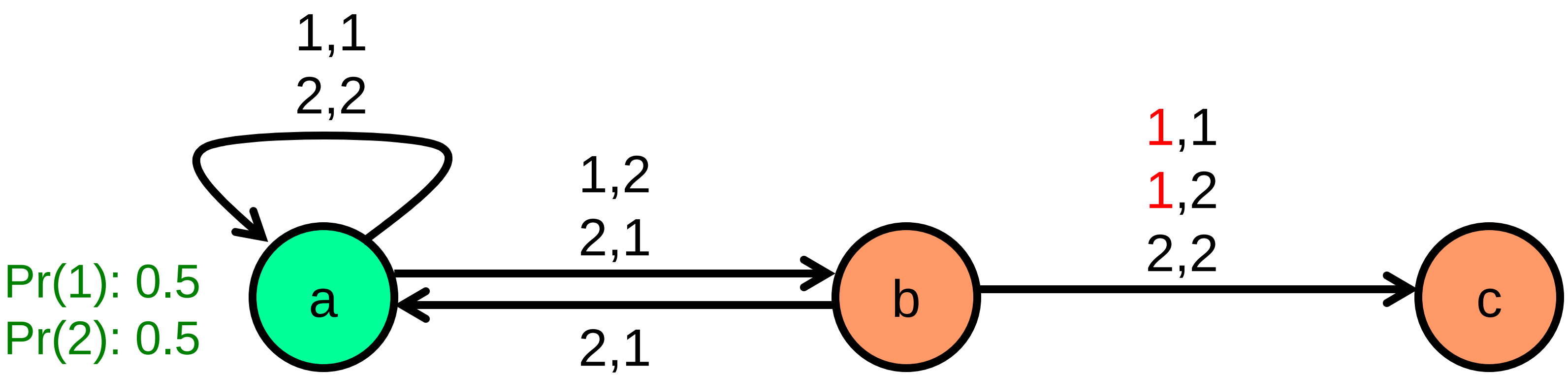
-for modeling open systems [Abramsky, Alur, Kupferman, Vardi, …]
-for strategy synthesis ("control") [Ramadge, Wonham, Pnueli, Rosner]

Property ◇c



⟨⟨left⟩⟩ ◇c … player "left" has a strategy to enforce ◇c

Property ◇c



⟨⟨left⟩⟩ ◇c … player "left" has a strategy to enforce ◇c

⟨left⟩ ◇c … player “left" has a randomized strategy to enforce ◇c

Qualitative Models

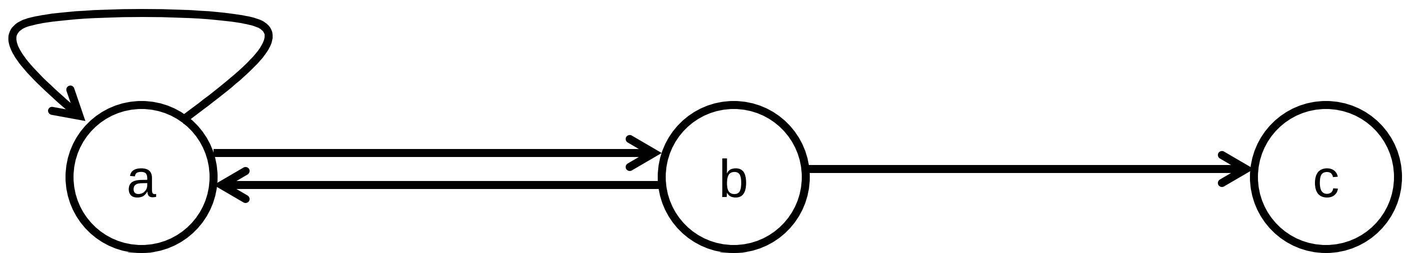
Trace: sequence of observations

Property p: assigns a reward to each trace
boolean rewards

Model m: generates a set of traces
(game) graph

Value(p,m): defined from the rewards of the generated traces
$\mathbb{B}$
$\exists$ or $\forall$ ($\exists\forall$)

Stochastic Game

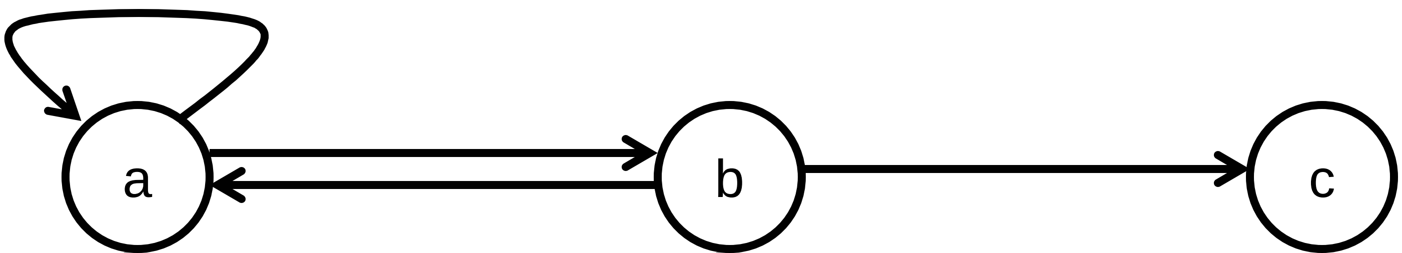


right / left	1	2
1	a: 0.6 b: 0.4	a: 0.5 b: 0.5
2	a: 0.1 b: 0.9	a: 0.2 b: 0.8

right / left	1	2
1	a: 0.0 c: 1.0	a: 0.0 c: 1.0
2	a: 0.7 b: 0.3	a: 0.0 b: 1.0

Property ◇c

Probability with which player "left" can enforce ◇c ?



State a:

left \ right	1	2
1	a: 0.6 b: 0.4	a: 0.5 b: 0.5
2	a: 0.1 b: 0.9	a: 0.2 b: 0.8

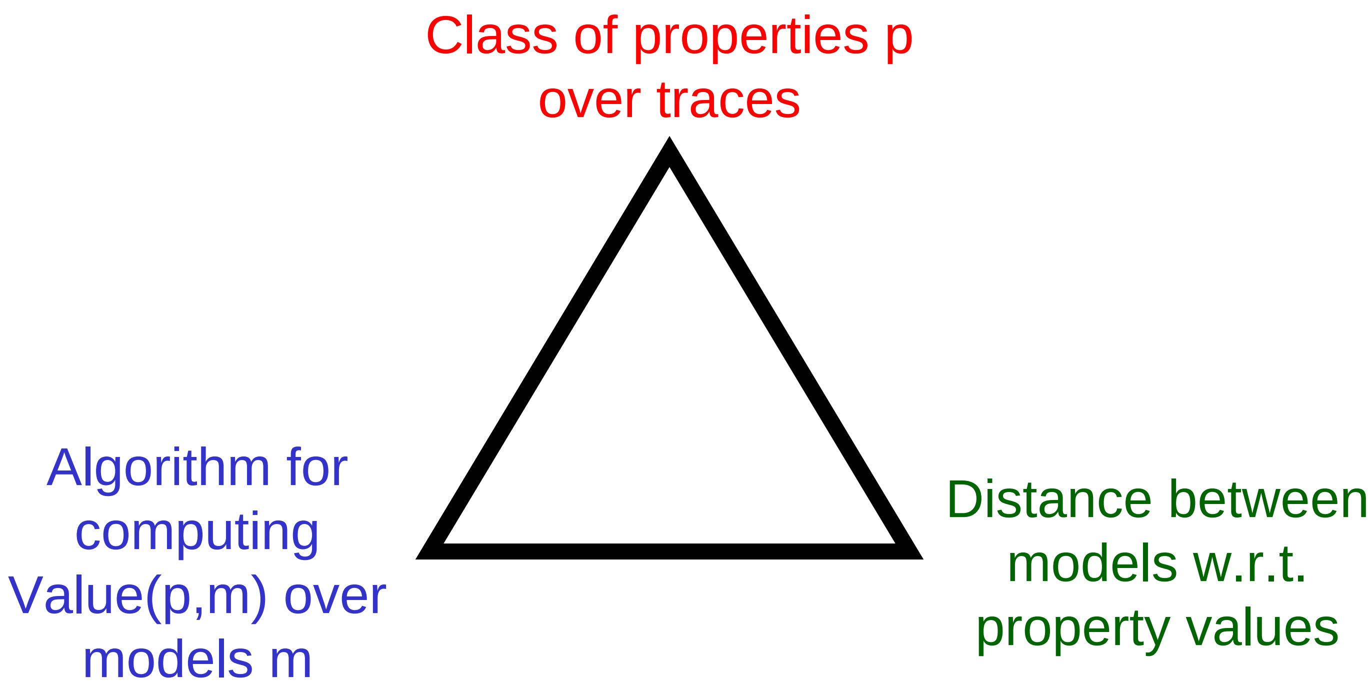
State b:

left \ right	1	2
1	a: 0.0 c: 1.0	a: 0.0 c: 1.0
2	a: 0.7 b: 0.3	a: 0.0 b: 1.0

Semi-Quantitative Models

Trace:	sequence of observations
Property p:	assigns a reward to each trace boolean rewards
Model m:	generates a set of traces (game) graph
Value(p,m): $[0,1] \subseteq \mathbb{R}$	defined from the rewards of the generated traces sup or inf (sup inf)

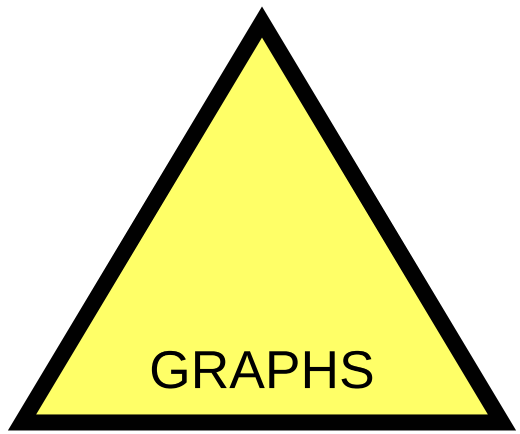
A Systems Theory



A Systems Theory

ω-regular properties

Class of properties p
over traces



Algorithm for
computing
Value(p,m) over
models m

μ-calculus

Distance between
models w.r.t.
property values

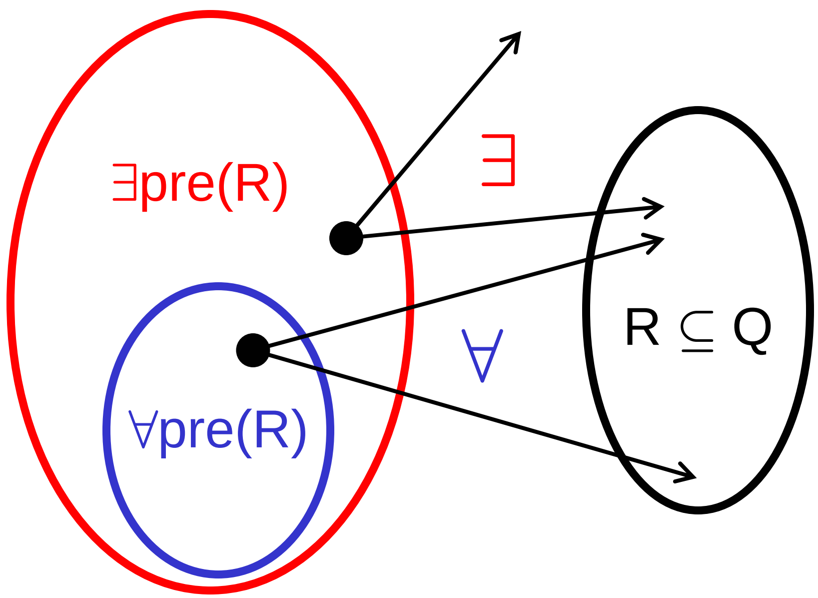
bisimilarity

Transition Graph

Q	states
δ: Q $\to 2^Q$	transition relation

Graph Regions

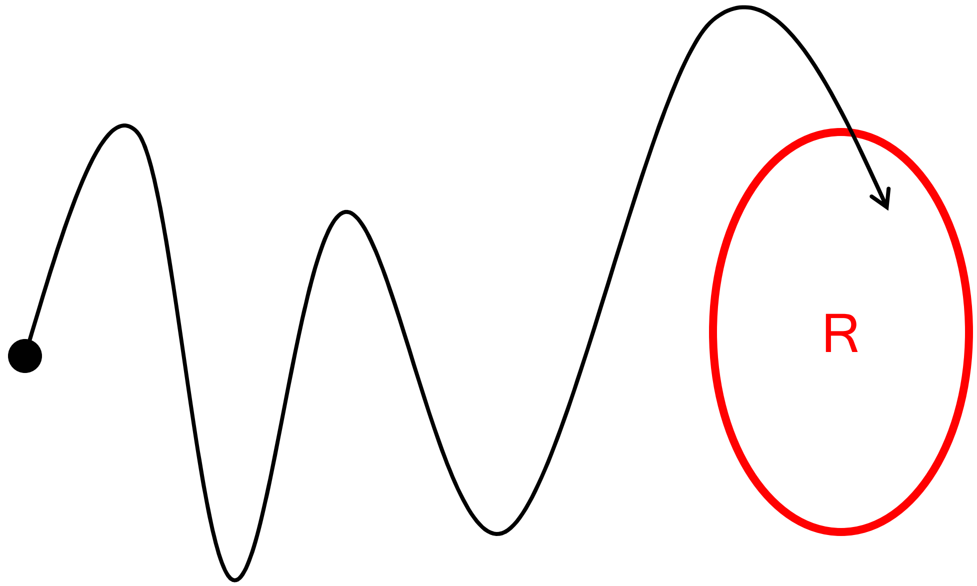
Q	states
δ: $Q \rightarrow 2^Q$	transition relation
$\Re = [\, Q \rightarrow \mathbb{B}\,]$	regions
$\exists$pre, $\forall$pre: $\Re \rightarrow \Re$	



Graph Property Values: Reachability

$\exists \Diamond R$

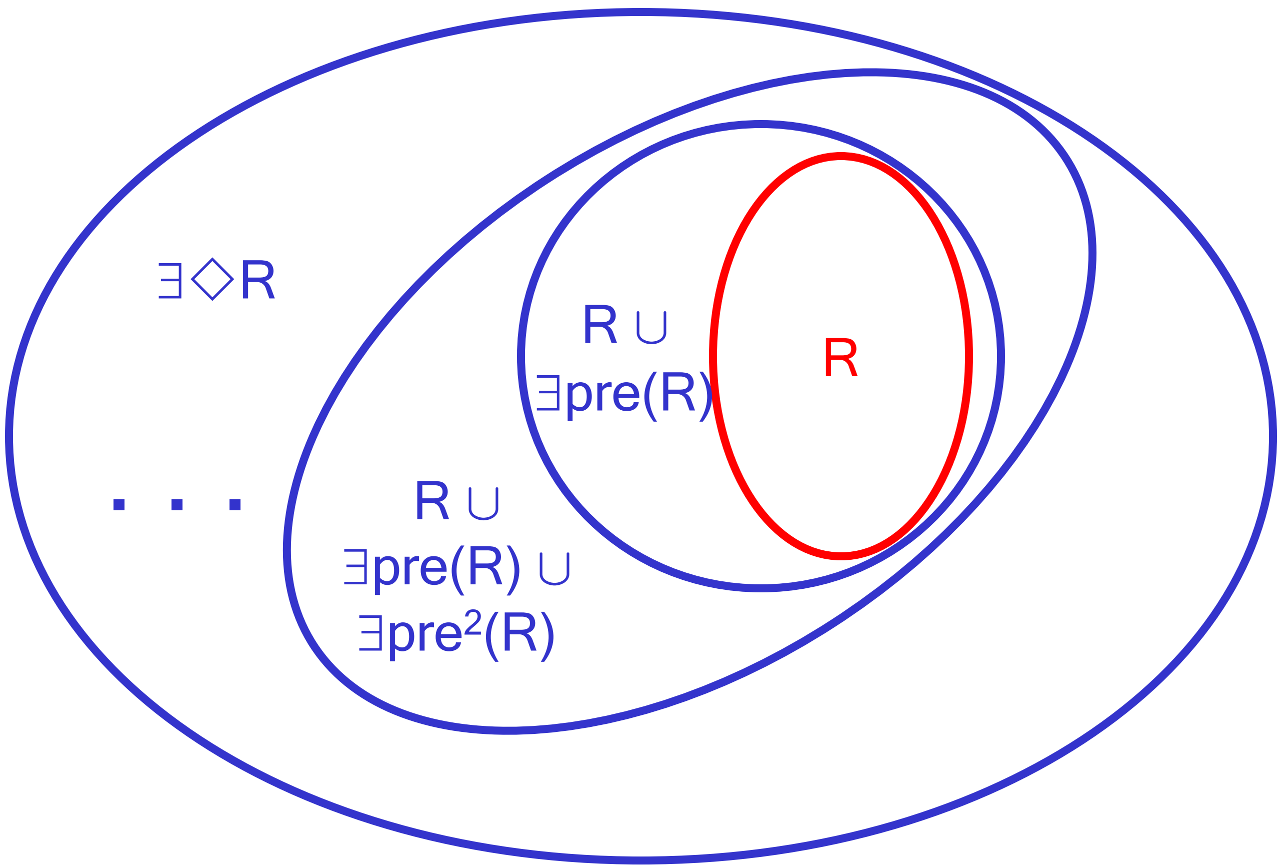
Given $R \subseteq Q$, find the states from which some trace leads to R.



Graph Property Values: Reachability

$\exists \Diamond R = (\mu X)(R \vee \exists pre(X))$

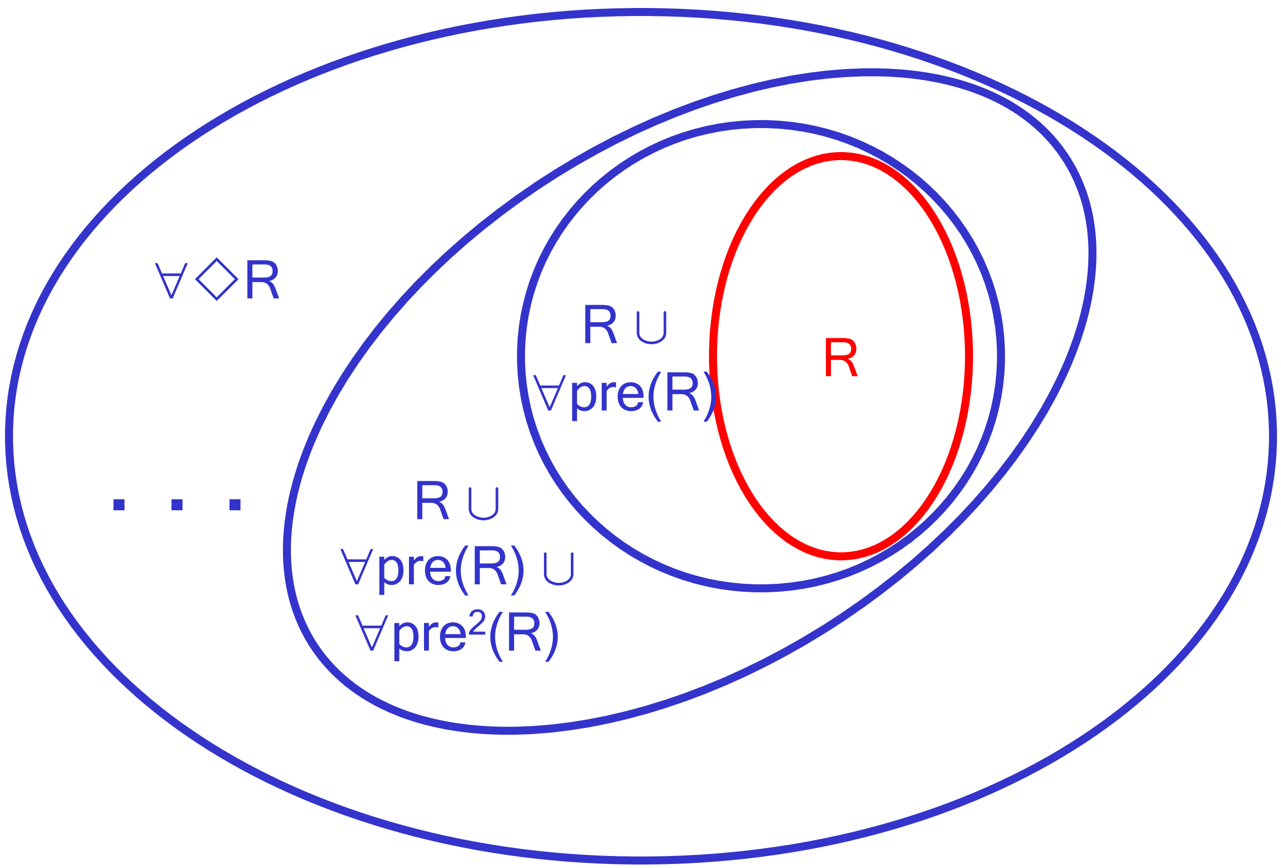
Given R⊆Q, find the states from which some trace leads to R.



Graph Property Values: Reachability

$\forall \Diamond R \;=\; (\mu X)\,(R \vee \forall pre(X))$

Given R⊆Q, find the states from which all traces lead to R.



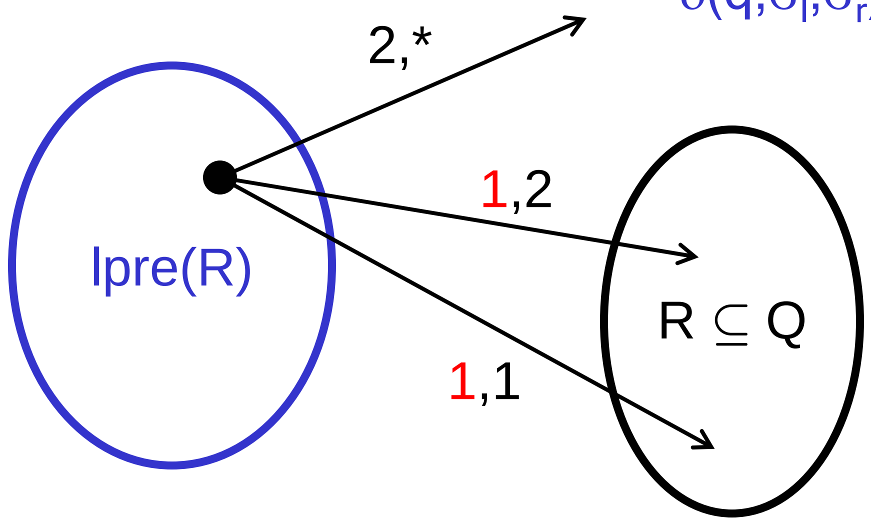
Concurrent Game

Q	states
Σ_l, Σ_r	moves of both players
δ: $Q \times \Sigma_l \times \Sigma_r \rightarrow Q$	transition function

Game Regions

Q	states
Σ_l, Σ_r	moves of both players
δ: $Q \times \Sigma_l \times \Sigma_r \to Q$	transition function
$\Re = [\, Q \to \mathbb{B} \,]$	regions
lpre, rpre: $\Re \to \Re$	

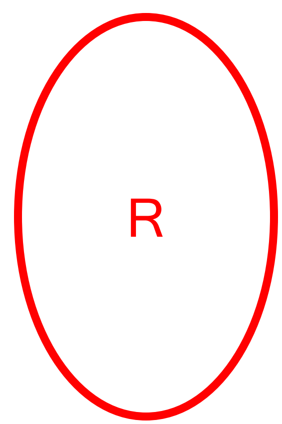
$q \in$ lpre(R) iff $(\exists \sigma_l \in \Sigma_l)\ (\forall \sigma_r \in \Sigma_r)\ \delta(q,\sigma_l,\sigma_r) \in R$



Game Property Values: Reachability

$\langle\langle \text{left} \rangle\rangle \Diamond R$

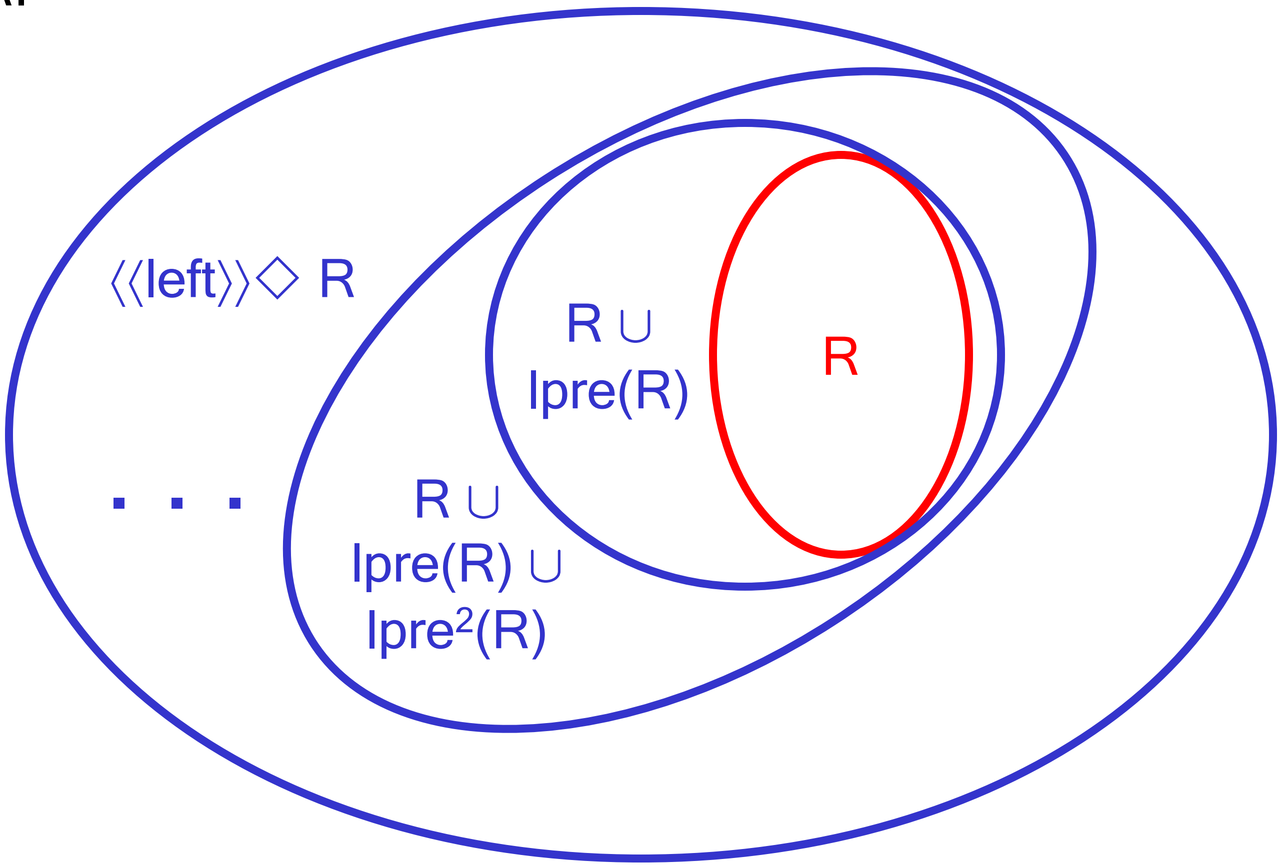
Given $R \subseteq Q$, find the states from which player "left" has a strategy to force the game to R.



Game Property Values: Reachability

$\langle\langle \text{left} \rangle\rangle \Diamond R \;=\; (\mu X)(R \vee \text{lpre}(X))$

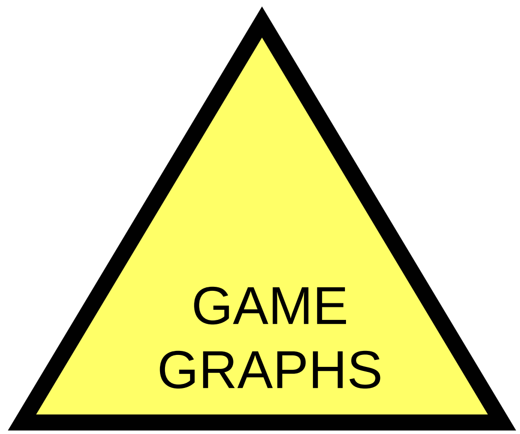
Given $R \subseteq Q$, find the states from which player "left" has a strategy to force the game to R.



An Open Systems Theory

ω-regular properties

Class of winning
conditions p over traces



Algorithm for
computing
Value(p,m) over
models m

(lpre,rpre) fixpoint calculus

Distance between
models w.r.t.
property values

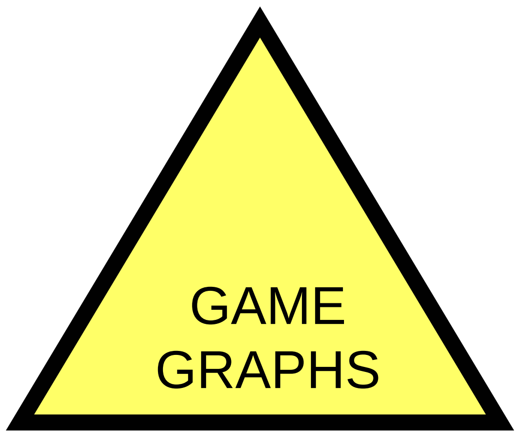
alternating bisimilarity

[Alur, H, Kupferman, Vardi]

An Open Systems Theory

ω-regular properties $\langle\langle left \rangle\rangle \Diamond R$

Class of winning conditions p over traces



Algorithm for computing Value(p,m) over models m

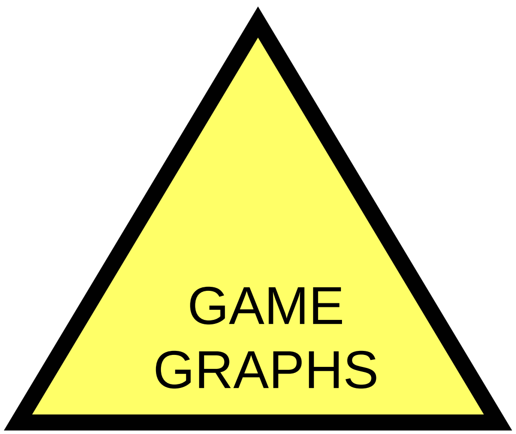
(lpre,rpre) fixpoint calculus

$(\mu X)\ (R \vee lpre(X))$

Every deterministic fixpoint formula ϕ computes Value(p,m), where p is the linear interpretation [Vardi] of ϕ.

An Open Systems Theory

Two states agree on the values of all fixpoint formulas iff they are alternating bisimilar [Alur, H, Kupferman, Vardi].



Algorithm for computing Value(p,m) over models m

Distance between models w.r.t. property values

(lpre,rpre) fixpoint calculus ↔ **alternating bisimilarity**

Stochastic Game

Q	states
Σ_l, Σ_r	moves of both players
$\delta: Q \times \Sigma_l \times \Sigma_r \to \mathrm{Dist}(Q)$	probabilistic transition function

Quantitative Game Regions

Q	states
Σ_l, Σ_r	moves of both players
δ: $Q \times \Sigma_l \times \Sigma_r \rightarrow \text{Dist}(Q)$	probabilistic transition function
$\Re = [\, Q \rightarrow [0,1] \,]$	quantitative regions

lpre, rpre: $\Re \rightarrow \Re$

$$\text{lpre}(R)(q) \quad = \quad (\sup \sigma_l \in \Sigma_l)\ (\inf \sigma_r \in \Sigma_r)\ R(\delta(q,\sigma_l,\sigma_r))$$

Quantitative Game Regions

Q	states
Σ_l, Σ_r	moves of both players
$\delta: Q \times \Sigma_l \times \Sigma_r \to Dist(Q)$	probabilistic transition function
$\mathfrak{R} = [\, Q \to [0,1] \,]$	quantitative regions

$\mathbb{B}$

lpre, rpre: $\mathfrak{R} \to \mathfrak{R}$

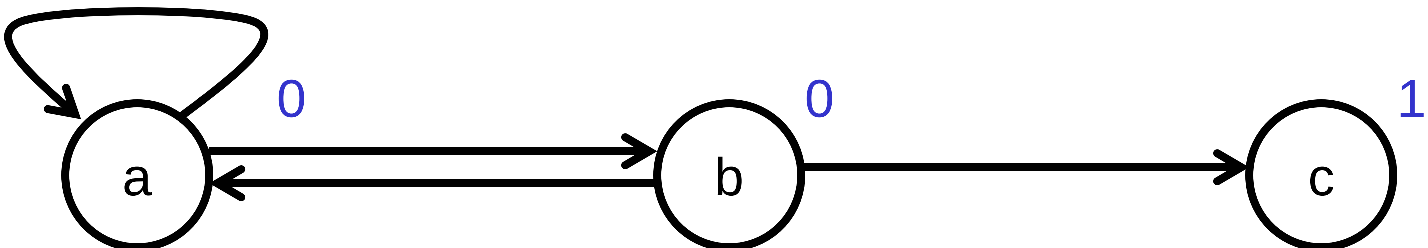
$$lpre(R)(q) = (\sup \sigma_l \in \Sigma_l)\ (\inf \sigma_r \in \Sigma_r)\ R(\delta(q,\sigma_l,\sigma_r))$$

$\exists$ $\forall$

Probability with which player "left" can enforce ◇c :

$(\mu X)\ (c \vee \mathrm{lpre}(X))$

$\vee$ = pointwise max



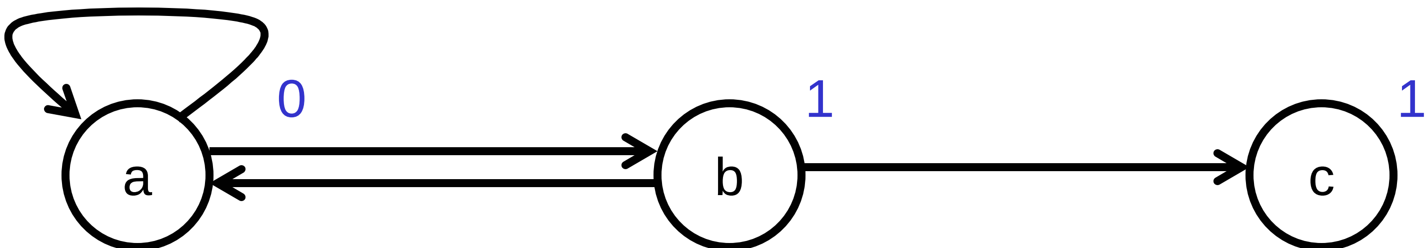
left \ right	1	2
1	**a: 0.6 b: 0.4**	**a: 0.5 b: 0.5**
2	**a: 0.1 b: 0.9**	**a: 0.2 b: 0.8**

left \ right	1	2
1	**a: 0.0 c: 1.0**	**a: 0.0 c: 1.0**
2	**a: 0.7 b: 0.3**	**a: 0.0 b: 1.0**

Probability with which player "left" can enforce $\Diamond c$:

$(\mu\ X)\ (c \vee \text{lpre}(X))$

$\vee$ = pointwise max



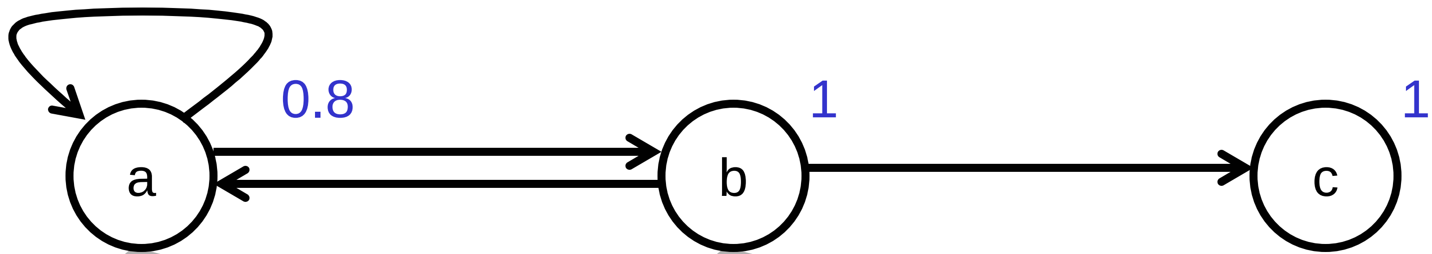
right / left	1	2
1	**a: 0.6** **b: 0.4**	**a: 0.5** **b: 0.5**
2	**a: 0.1** **b: 0.9**	**a: 0.2** **b: 0.8**

right / left	1	2
1	**a: 0.0** **c: 1.0**	**a: 0.0** **c: 1.0**
2	**a: 0.7** **b: 0.3**	**a: 0.0** **b: 1.0**

Probability with which player "left" can enforce ◇c :

$(\mu X)\ (c \vee lpre(X))$

$\vee$ = pointwise max



State a:

left \ right	1	2
1	**a: 0.6 b: 0.4**	**a: 0.5 b: 0.5**
2	**a: 0.1 b: 0.9**	**a: 0.2 b: 0.8**

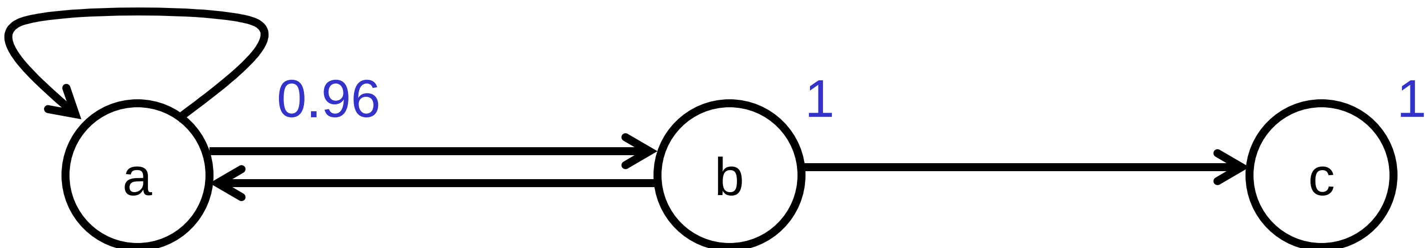
State b:

left \ right	1	2
1	**a: 0.0 c: 1.0**	**a: 0.0 c: 1.0**
2	**a: 0.7 b: 0.3**	**a: 0.0 b: 1.0**

Probability with which player "left" can enforce ◇c :

$(\mu X)(c \vee lpre(X))$

$\vee$ = pointwise max



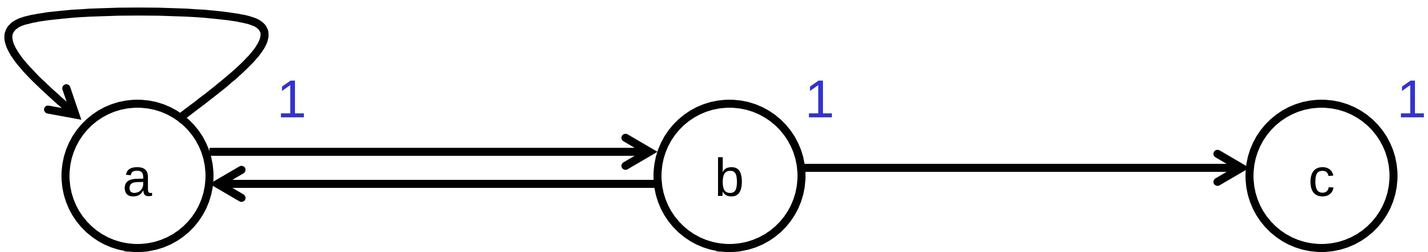
right / left	1	2
1	**a: 0.6** **b: 0.4**	**a: 0.5** **b: 0.5**
2	**a: 0.1** **b: 0.9**	**a: 0.2** **b: 0.8**

right / left	1	2
1	**a: 0.0** **c: 1.0**	**a: 0.0** **c: 1.0**
2	**a: 0.7** **b: 0.3**	**a: 0.0** **b: 1.0**

Probability with which player "left" can enforce ◇c :

$(\mu X)(c \vee lpre(X))$

$\vee$ = pointwise max



left \ right	1	2
1	**a: 0.6** **b: 0.4**	**a: 0.5** **b: 0.5**
2	**a: 0.1** **b: 0.9**	**a: 0.2** **b: 0.8**

left \ right	1	2
1	**a: 0.0** **c: 1.0**	**a: 0.0** **c: 1.0**
2	**a: 0.7** **b: 0.3**	**a: 0.0** **b: 1.0**

In the limit, the deterministic fixpoint formulas work for all ω-regular properties [de Alfaro, Majumdar].

A Probabilistic Systems Theory

ω-regular properties

Class of properties p
over traces

Algorithm for
computing
Value(p,m) over
models m

Distance between
models w.r.t.
property values

quantitative fixpoint calculus

quantitative bisimilarity
[Desharnais, Gupta, Jagadeesan, Panangaden]

A Probabilistic Systems Theory

quantitative ω-regular properties

Class of properties p
over traces

max expected value
of satisfying ◇R

Algorithm for
computing
Value(p,m) over
models m

quantitative fixpoint calculus

$(\mu X)\ (R \vee \exists pre(X))$

Every deterministic
fixpoint formula ϕ
computes expected
Value(p,m), where p
is the linear
interpretation of ϕ.

Qualitative Bisimilarity

$e: Q^2 \to \{0,1\}$ … equivalence relation

F … function on equivalences

$F(e)(q,q') = 0$ if q and q' disagree on observations

$= \min \{ e(r,r') \mid r \in \exists pre(q) \wedge r' \in \exists pre(q') \}$ else

Qualitative bisimilarity … greatest fixpoint of F

Quantitative Bisimilarity

$d: Q^2 \to [0,1]$ … pseudo-metric ("distance")

F … function on pseudo-metrics

$$F(d)(q,q') = 1 \quad \text{if } q \text{ and } q' \text{ disagree on observations}$$
$$\approx \text{max of } \quad \sup_l \inf_r \ d(\delta(q,l,r),\delta(q',l,r))$$
$$\sup_r \inf_l \ d(\delta(q,l,r),\delta(q',l,r)) \quad \text{else}$$

Quantitative bisimilarity … greatest fixpoint of F

Natural generalization of bisimilarity from binary relations to pseudo-metrics.

A Probabilistic Systems Theory

Two states agree on the values of all quantitative fixpoint formulas iff their quantitative bisimilarity distance is 0.

MARKOV DECISION PROCESSES

Algorithm for computing Value(p,m) over models m

Distance between models w.r.t. property values

quantitative fixpoint calculus ↔ **quantitative bisimilarity**

Great – BUT …

1 The theory is too precise.

Even the smallest change in the probability of a transition can cause an arbitrarily large change in the value of a property.

2 The theory is not computational.

We cannot bound the rate of convergence for quantitative fixpoint formulas.

Solution: Discounting

Economics:

A dollar today is better than a dollar tomorrow.

Value of $1.- today:	1	
Tomorrow:	α	for discount factor $0 < \alpha < 1$
Day after tomorrow:	α^2	
etc.		

Solution: Discounting

Economics:

A dollar today is better than a dollar tomorrow.

Value of $1.- today:	1	
Tomorrow:	α	for discount factor $0 < \alpha < 1$
Day after tomorrow:	α^2	
etc.		

Engineering:

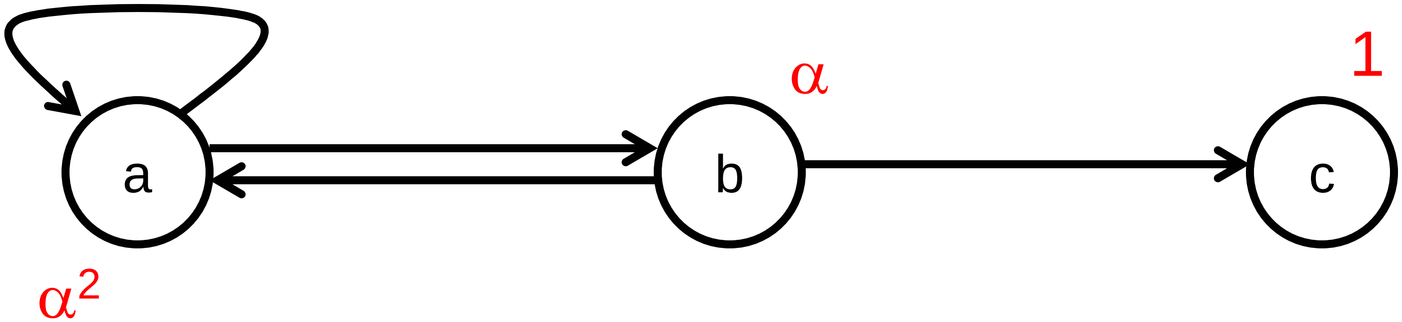
A bug today is worse than a bug tomorrow.

Discounted Reachability

$$\text{Reward}\,(\Diamond_{\alpha}\, c\,) = \begin{cases} \alpha^{k} & \text{if} \quad c \text{ is first true after } k \text{ transitions} \\ 0 & \text{if} \quad c \text{ is never true} \end{cases}$$

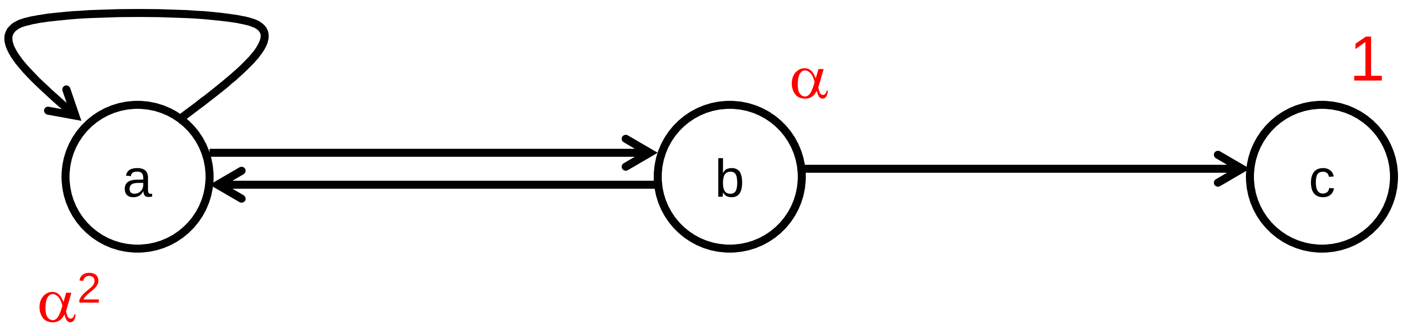
The reward is proportional to how quickly c is satisfied.

Discounted Property $\Diamond_\alpha$ c



$\exists \Diamond_\alpha c$

Discounted Property $\Diamond_\alpha c$



$\exists \Diamond_\alpha c$

Discounted fixpoint calculus: $\text{pre}(\phi) \longrightarrow \alpha \cdot \text{pre}(\phi)$

Fully Quantitative Models

Trace:	sequence of observations
Property p:	assigns a reward to each trace real reward
Model m:	generates a set of traces (game) graph
Value(p,m): [0,1] $\subseteq \mathbb{R}$	defined from the rewards of the generated traces sup or inf (sup inf)

Discounted Bisimilarity

$d: Q^2 \rightarrow [0,1]$ … pseudo-metric ("distance")

F … function on pseudo-metrics

$F(d)(q,q') = 1$ if q and q' disagree on observations

$\approx$ max of $\sup_l \inf_r d(\delta(q,l,r),\delta(q',l,r))$

$\sup_r \inf_l d(\delta(q,l,r),\delta(q',l,r))$ else

Quantitative bisimilarity … greatest fixpoint of F

$\alpha \cdot$

A Discounted Systems Theory

discounted ω-regular properties

Class of winning
rewards p over traces

Algorithm for
computing
Value(p,m) over
models m

Distance between
models w.r.t.
property values

discounted fixpoint calculus

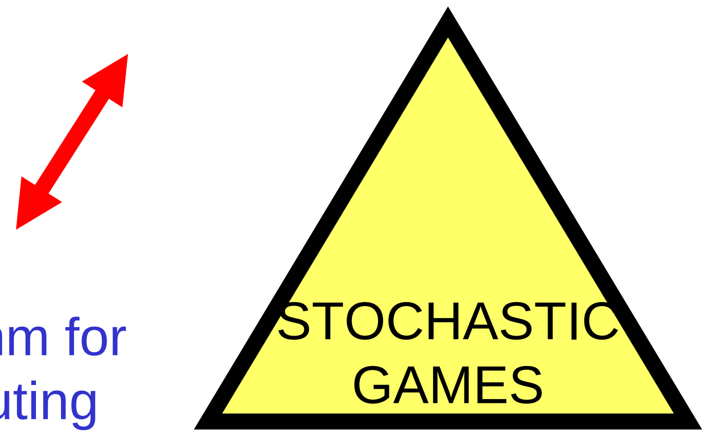
discounted bisimilarity

A Discounted Systems Theory

discounted ω-regular properties

Class of expected rewards p over traces

max expected reward $\Diamond_{\alpha} R$ achievable by left player



Algorithm for computing Value(p,m) over models m

Every discounted deterministic fixpoint formula ϕ computes Value(p,m), where p is the linear interpretation of ϕ.

discounted fixpoint calculus

$(\mu X)\ (R \vee \alpha \cdot \mathrm{lpre}(X))$

A Discounted Systems Theory

The difference between two states in the values of discounted fixpoint formulas is bounded by their discounted bisimilarity distance.

Algorithm for computing Value(p,m) over models m

Distance between models w.r.t. property values

discounted fixpoint calculus ⟷ **discounted bisimilarity**

Discounting is Robust

Continuity over Traces:

Every discounted fixpoint formula defines a reward function on traces that is continuous in the Cantor metric.

Continuity over Models:

If transition probabilities are perturbed by ε, then discounted bisimilarity distances change by at most $f(\varepsilon)$.

Discounting is robust against effects at infinity, and against numerical perturbations.

Discounting is Computational

The iterative evaluation of an α-discounted fixpoint formula converges geometrically in α.

(So we can compute to any desired precision.)

Discounting is Approximation

If the discount factor tends towards 1, then we recover the classical theory:

- $\lim_{\alpha \to 1}$ α-discounted interpretation of fixpoint formula ϕ
  = classical interpretation of ϕ
- $\lim_{\alpha \to 1}$ α-discounted bisimilarity
  = classical (alternating; quantitative) bisimilarity

Further Work

- Exact computation of discounted values of temporal formulas over finite-state systems [de Alfaro, Faella, H, Majumdar, Stoelinga].
- Discounting real-time systems: continuous discounting of time delay rather than discrete discounting of number of steps [Prabhu].

Conclusions

- Discounting provides a continuous and computational approximation theory of discrete and probabilistic processes.
- Discounting captures an important engineering intuition.

"In the long run, we're all dead." J.M. Keynes