

---

# Probabilistic Automata: Probability and Nondeterminism Together

Roberto Segala  
University of Verona



# Objectives

- Why/how probability and nondeterminism?
- What are the key ingredients?
  - Measures and measurable functions
  - Conservative extensions of Labeled Transition Systems
- Probabilistic Automata
  - Definition
  - Comparison with other models
  - Parallel composition
- Bisimulation relations
  - Strong, strong probabilistic, weak
- Language inclusion
- Highlights of a Case study



# Why Nondeterminism with Probability?

## Distributed Algorithms

---

- Some problems are unsolvable
  - Consensus [FLP85]
- ... but are solvable with randomization
  - Probabilistic consensus [Ben83,AH90]
- Probability and nondeterminism coexist
  - Probability:
    - Processes flip coins
  - Nondeterminism:
    - Several processes in parallel
    - Do not care whether the coin is fair
- Quantitative analysis
  - What is the worst expected complexity?



# Why Nondeterminism with Probability?

## Stochastic Games

---

- Nondeterminism
  - Each player has several moves available
- Probability
  - Moves may involve coin flipping
- Quantitative analysis
  - What is the best probability to win the game?



# Why Nondeterminism with Probability?

## Security

---

- Nondeterminism
  - User behavior (adversary in Dolev-Yao)
  - Relative speeds of agents
  - Agent behavior (usually deterministic)
- Probability
  - Users and agents flip coins
    - Nonces, keys, random protocols
- Quantitative analysis
  - Probability of attack (negligible)



# Why Nondeterminism with Probability?

## Concurrency Theory

---

- Nondeterminism
  - Scheduling within parallel composition
  - Unknown behavior of the environment
  - Underspecification
- Probability
  - Environment may be stochastic
  - Processes may flip coins



# Probability and Nondeterminism: How?

- **Reactive, Generative Systems [LS89,GSST90]**
  - Labeled transition systems
    - Add probabilities to the arcs
  - Process algebras
    - Replace + with probabilistic +
- **Probabilistic Automata [Seg95]**
  - Labeled transition systems
    - Replace target states with target measures in transitions
  - Process Algebras
    - Add a probabilistic + operator (named  $\oplus$ )



# Automata

$$A = (Q, q_0, E, H, D)$$

Transition relation

$$D \subseteq Q \times (E \cup H) \times Q$$

Internal (hidden) actions

External actions:  $E \cap H = \emptyset$

Initial state:  $q_0 \in Q$

States



# Probabilistic Automata

$$PA = (Q, q_0, E, H, D)$$

Transition relation

$$D \subseteq Q \times (E \cup H) \times \text{Disc}(Q)$$

Internal (hidden) actions

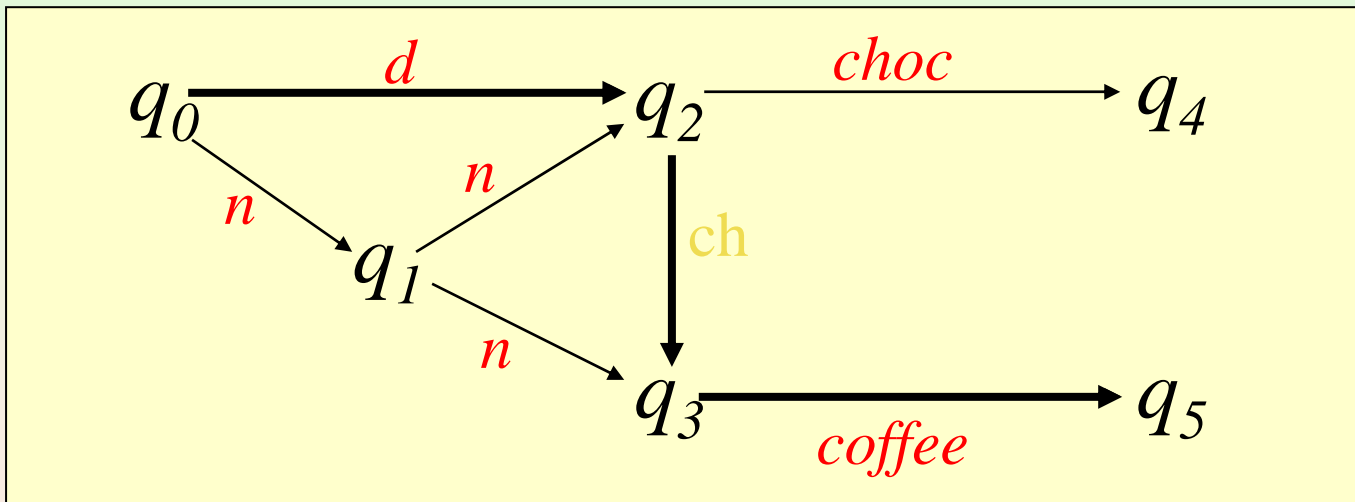
External actions:  $E \cap H = \emptyset$

Initial state:  $q_0 \in Q$

States

# Example: Automata

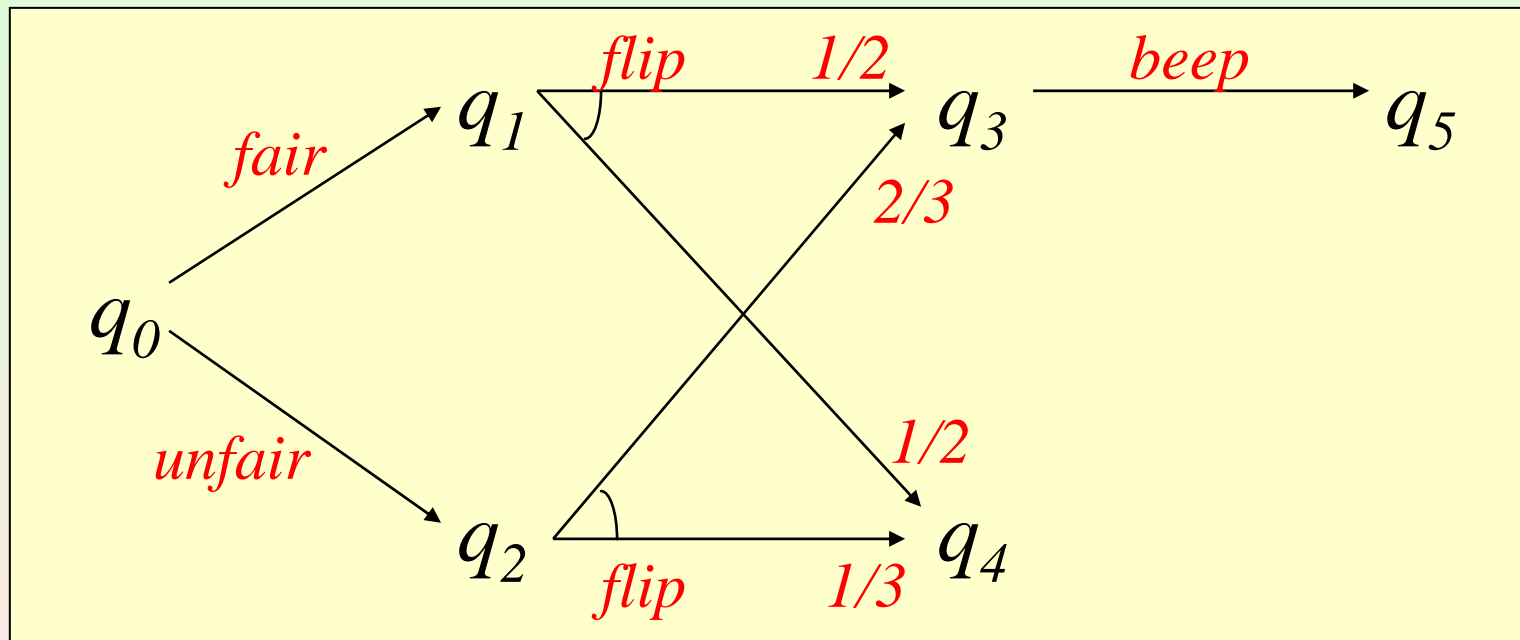
$$A = (Q, q_0, E, H, D)$$



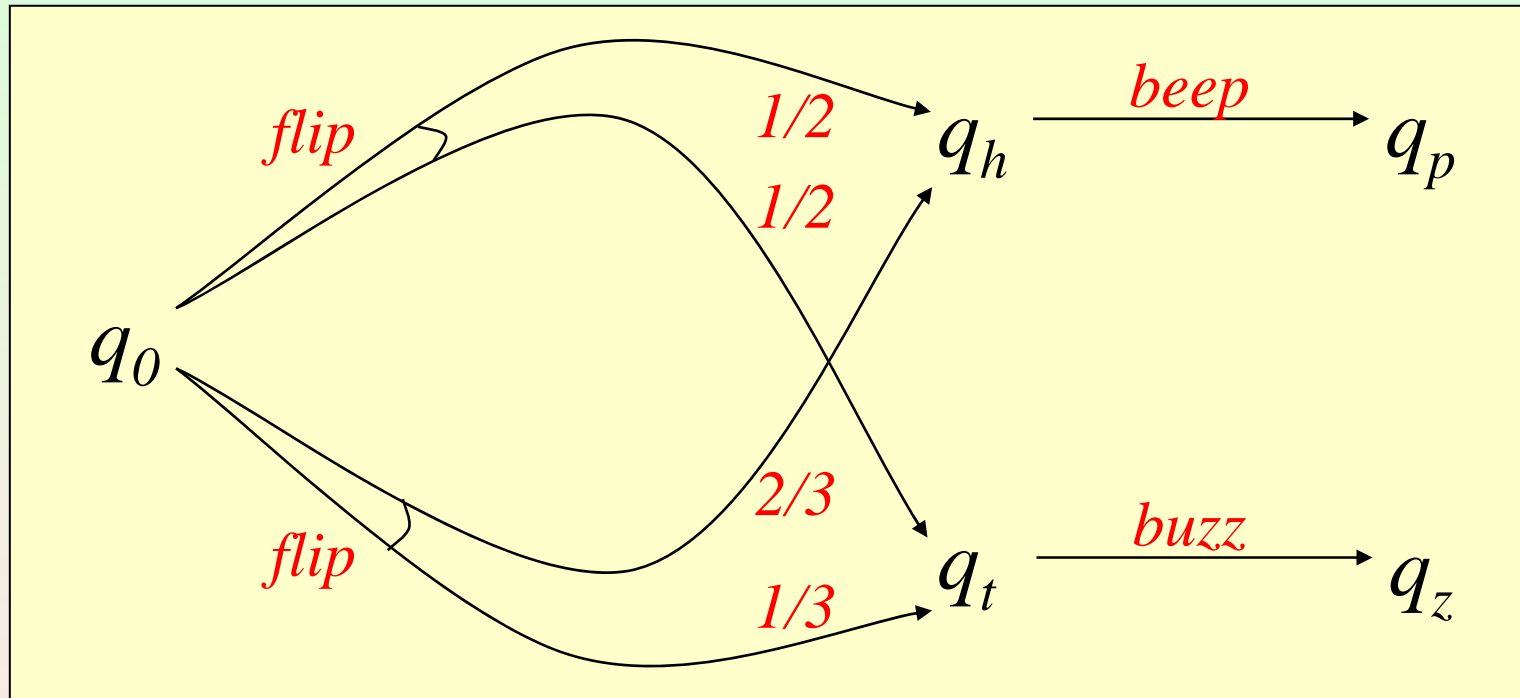
Execution:  $q_0 \xrightarrow{n} q_1 \xrightarrow{n} q_2 \xrightarrow{ch} q_3 \xrightarrow{coffee} q_5$

Trace:  $n \ n \ coffee$

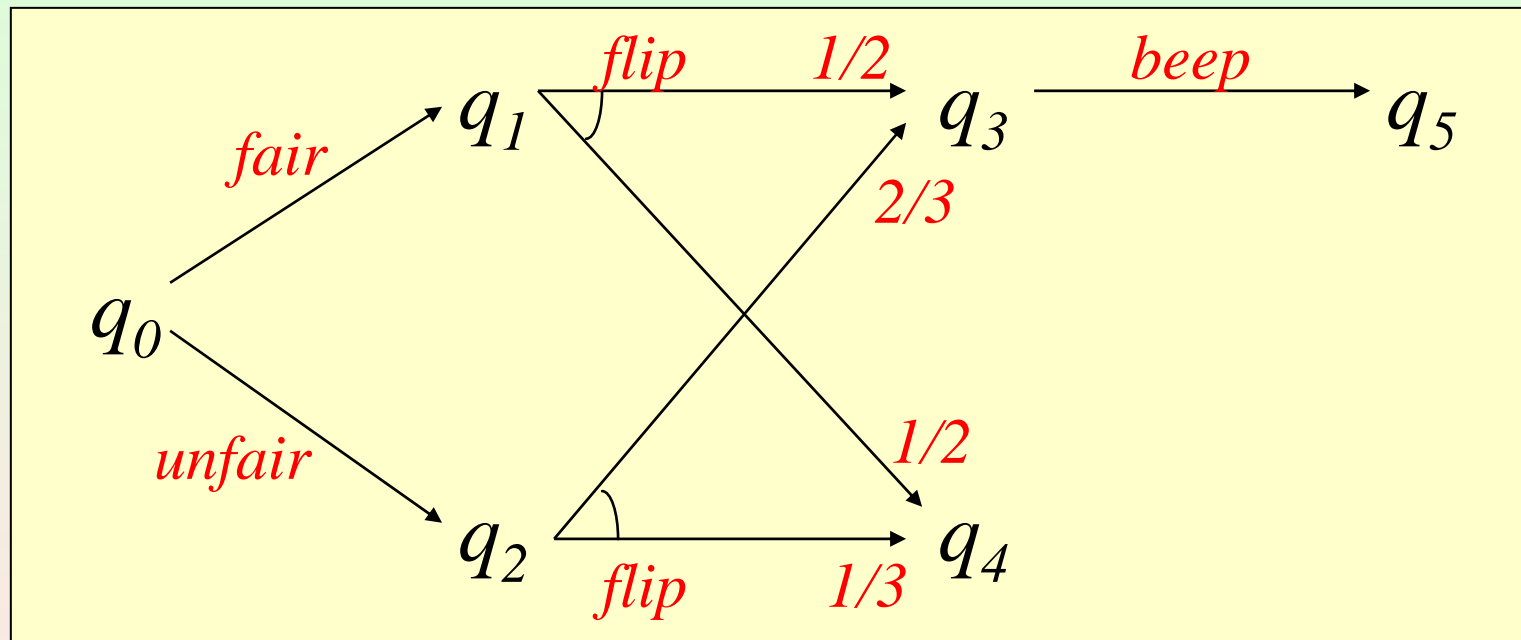
# Example: Probabilistic Automata



# Example: Probabilistic Automata

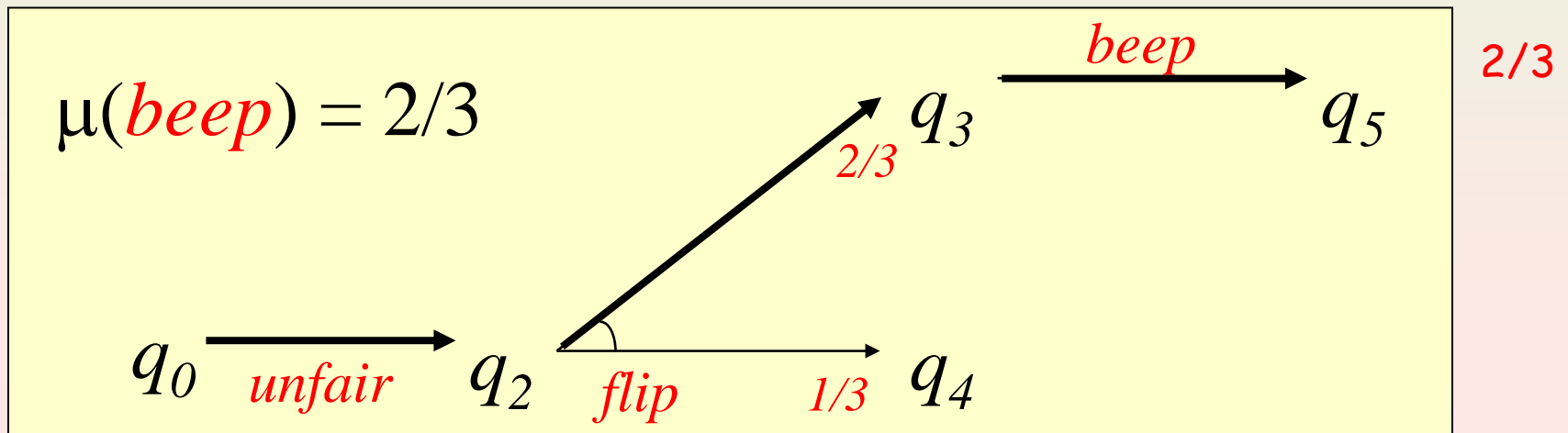
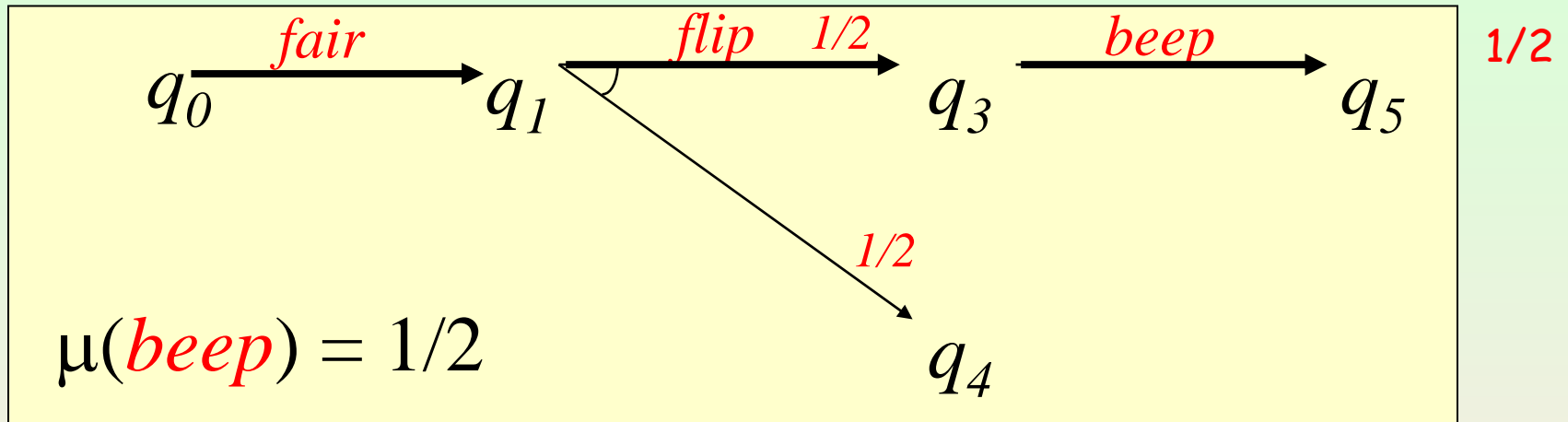


# Example: Probabilistic Automata

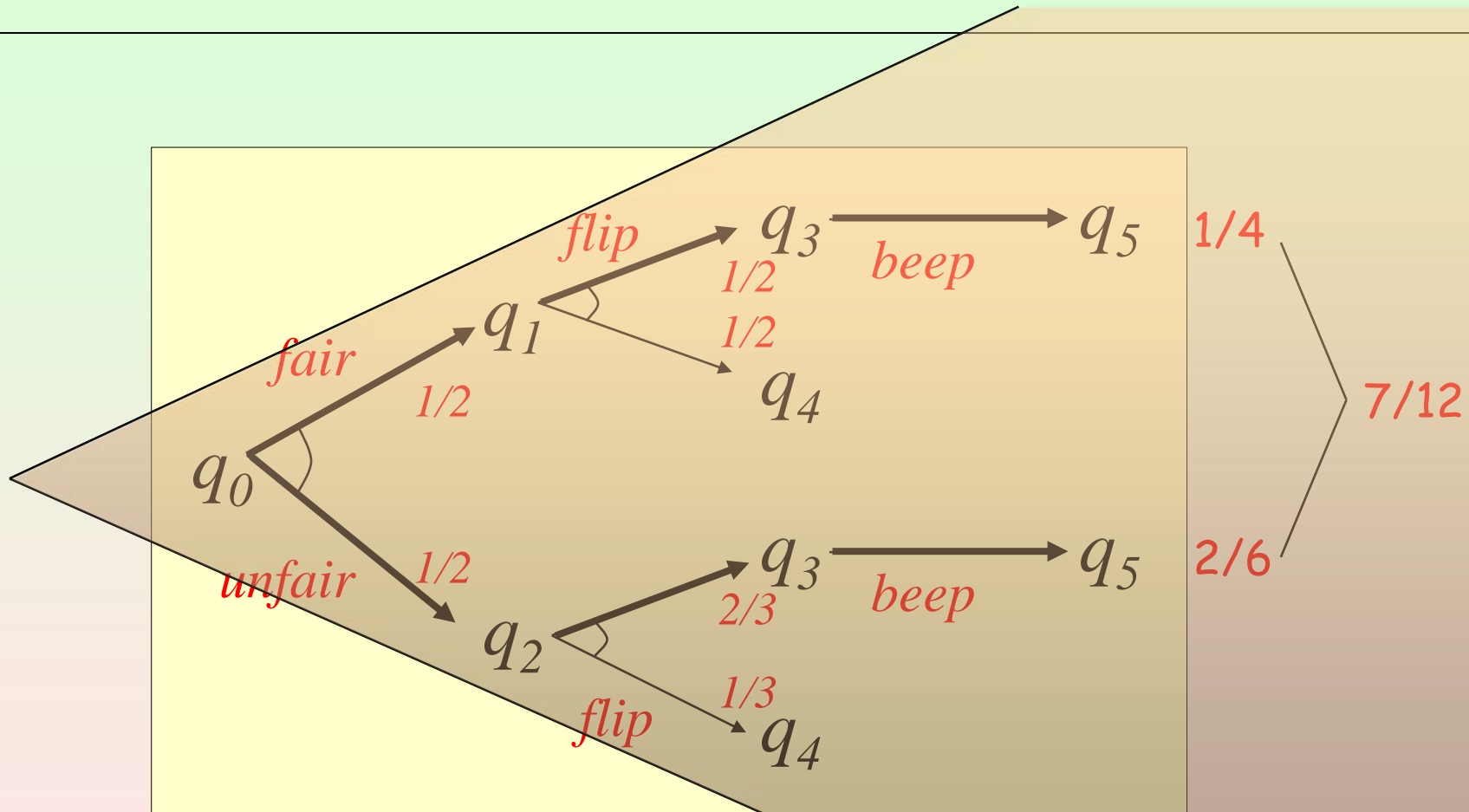


What is the probability of beeping?

# Example: Probabilistic Executions



# Example: Probabilistic Executions



# Measure Theory

- Sample set
  - Set of objects  $\Omega$
- Sigma-field ( $\sigma$ -field)
  - Subset  $F$  of  $2^\Omega$  satisfying
    - Inclusion of  $\Omega$
    - Closure under complement
    - Closure under countable union
    - Closure under countable intersection
- Measure on  $(\Omega, F)$ 
  - Function  $\mu$  from  $F$  to  $\mathcal{R}^{\geq 0}$ 
    - For each countable collection  $\{X_i\}_I$  of pairwise disjoint sets of  $F$ ,  $\mu(\cup_I X_i) = \sum_I \mu(X_i)$
- (Sub-)probability measure
  - Measure  $\mu$  such that  $\mu(\Omega) = 1$  ( $\mu(\Omega) \leq 1$ )
- Sigma-field generated by  $C \subseteq 2^\Omega$ 
  - Smallest  $\sigma$ -field that includes  $C$

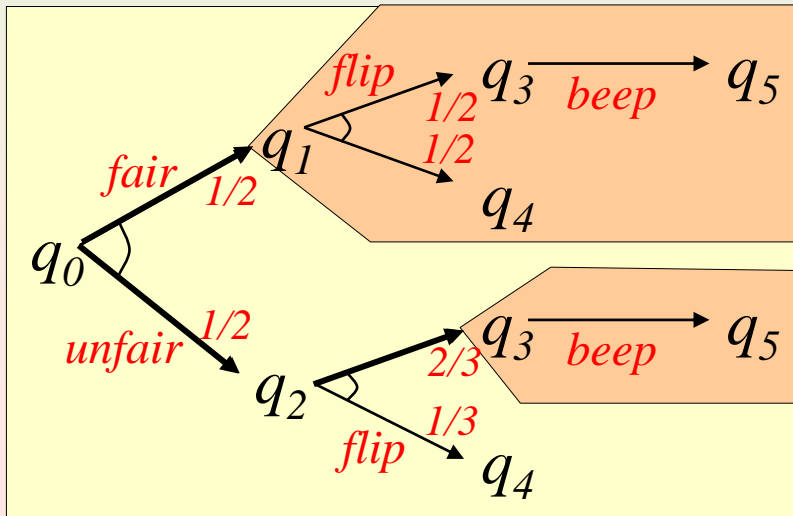
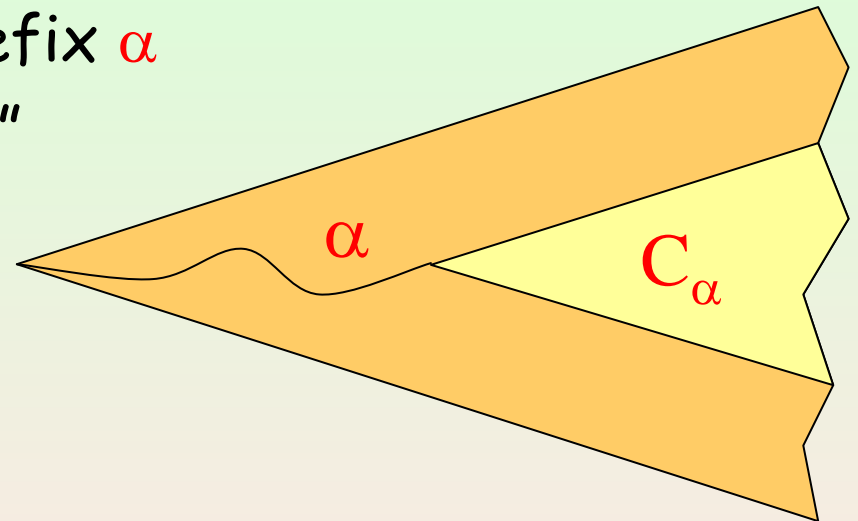
Why not  $F = 2^\Omega$ ?  
Example: set of executions  
Flip a fair coin infinitely many times  
 $\Omega = \{h, t\}^\infty$   
Study probabilities of sets of executions  
 $\mu(\omega) > 0$  for each  $\omega \in \Omega$   
 $\mu(\text{first coin } h) = 1/2$   
which sets can I measure?  
Theorem: there is no probability measure on  $2^\Omega$  such that  $\mu(\omega) = 0$  for each  $\omega \in \Omega$ .





# Cones and Measures

- Cone of  $\alpha$ 
  - Set of executions with prefix  $\alpha$
  - Represent event " $\alpha$  occurs"
- Measure of a cone
  - Product edges of  $\alpha$



extends uniquely  
 $\sigma$ -field generated by cones

# Examples of Events

- Eventually action  $a$  occurs
  - Union of cones where action  $a$  occurs once
- Action  $a$  occurs at least  $n$  times
  - Union of cones where action  $a$  occurs  $n$  times
- Action  $a$  occurs at most  $n$  times
  - Complement of **action  $a$  occurs at least  $n+1$  times**
- Action  $a$  occurs exactly  $n$  times
  - Intersection of previous two events
- Action  $a$  occurs infinitely many times
  - Intersection of **action  $a$  occurs at least  $n$  times** for all  $n$
- Execution  $\alpha$  occurs and nothing is scheduled after
  - Set consisting of  $\alpha$  only
  - $C_\alpha$  intersected complement of cones that extend  $\alpha$



# Schedulers - Probabilistic Executions

## Scheduler

Function

$$\sigma : \text{exec}^*(A) \rightarrow \text{SubDisc}(D)$$

if  $\sigma(\alpha)((q, a, v)) > 0$  then  $q = \text{lstate}(\alpha)$

## Probabilistic execution

generated by  $\sigma$  from state  $r$

Measure

$\mu_{\sigma, r}$

$$\mu_{\sigma, r}(C_s) = 0 \quad \text{if } r \neq s$$

$$\mu_{\sigma, r}(C_r) = 1$$

$$\mu_{\sigma, r}(C_{\alpha a q}) = \mu_{\sigma, r}(C_\alpha) \cdot \left( \sum_{(s, a, v) \in D} \sigma(\alpha)((s, a, v)) v(q) \right)$$

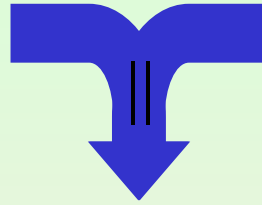
# Other Models

- Reactive and generative systems
  - Restricted forms of transitions
- Labeled Concurrent Markov Chains
  - Restricted forms of transitions
- Rabin's Probabilistic Automata
  - Introduced in the context of language theory
  - Extended by our Probabilistic Automata
- Unlabeled systems [Var85,BA95,BK98]
  - Can be Probabilistic Automata with a single invisible action
  - Labels may be associated with states
  - The theory does not change
- Markov Chains
  - Unlabeled systems that enable one transition from each state
- Probabilistic Input/Output Automata
  - Add Input/Output distinction on actions
  - Useful to handle composition of generative PAs



# Composition of Probabilistic Automata

$$A_1 = (Q_1, q_1, E_1, H_1, D_1)$$



$$A_2 = (Q_2, q_2, E_2, H_2, D_2)$$

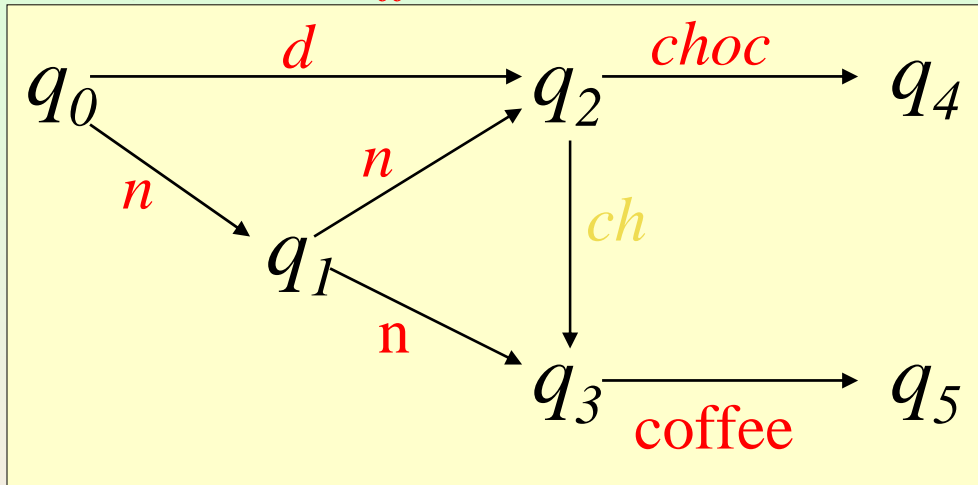
$$A_1 \parallel A_2 = (Q_1 \times Q_2, (q_1, q_2), E_1 \cup E_2, H_1 \cup H_2, D)$$

$$D = \left\{ (q, a, (s_1, s_2)) \mid \begin{array}{l} \text{if } a \in E_i \cup H_i \text{ then } (\pi_i(q), a, s_i) \in D_i \\ \text{if } a \notin E_i \cup H_i \text{ then } s_i = \pi_i(q) \end{array} \quad i \in \{1, 2\} \right\}$$

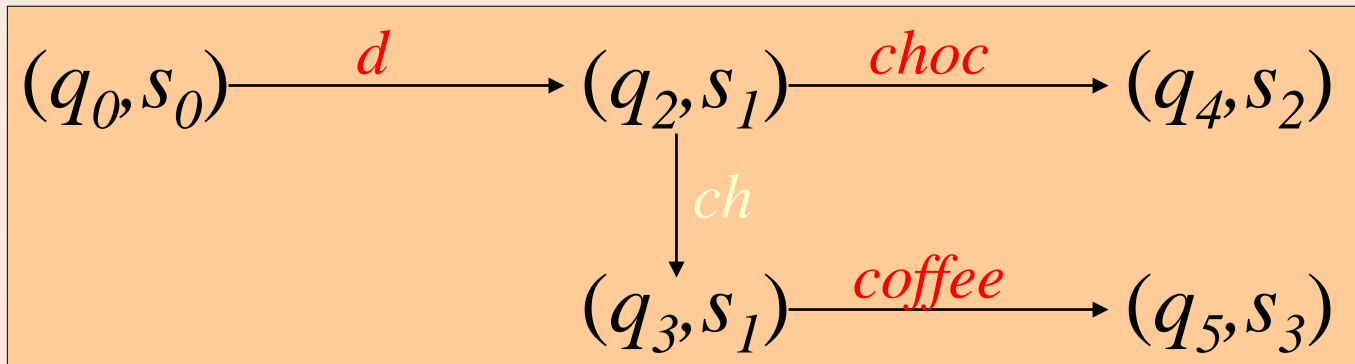
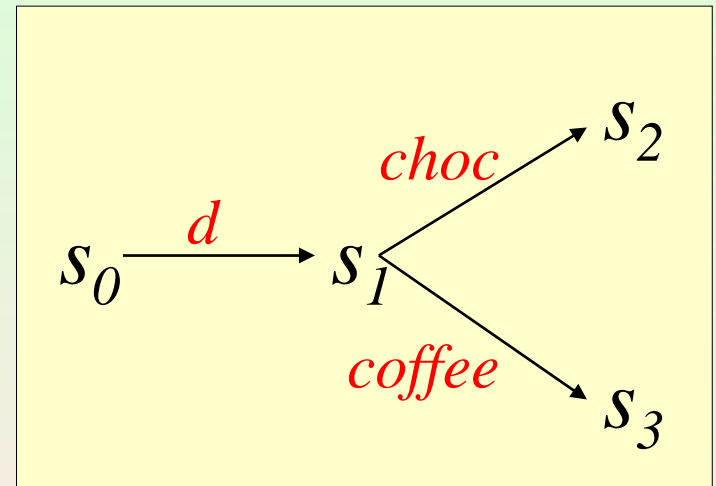
$$D = \left\{ (q, a, \mu_1 \times \mu_2) \mid \begin{array}{l} \text{if } a \in E_i \cup H_i \text{ then } (\pi_i(q), a, \mu_i) \in D_i \\ \text{if } a \notin E_i \cup H_i \text{ then } \mu_i = \delta(\pi_i(q)) \end{array} \quad i \in \{1, 2\} \right\}$$

# Example: Composition of Automata

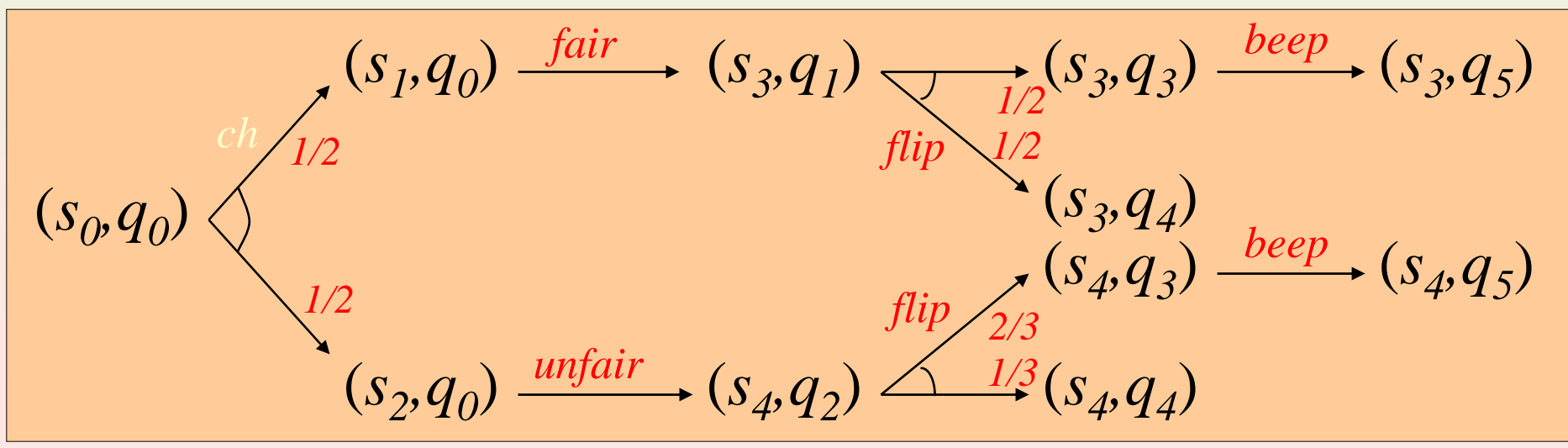
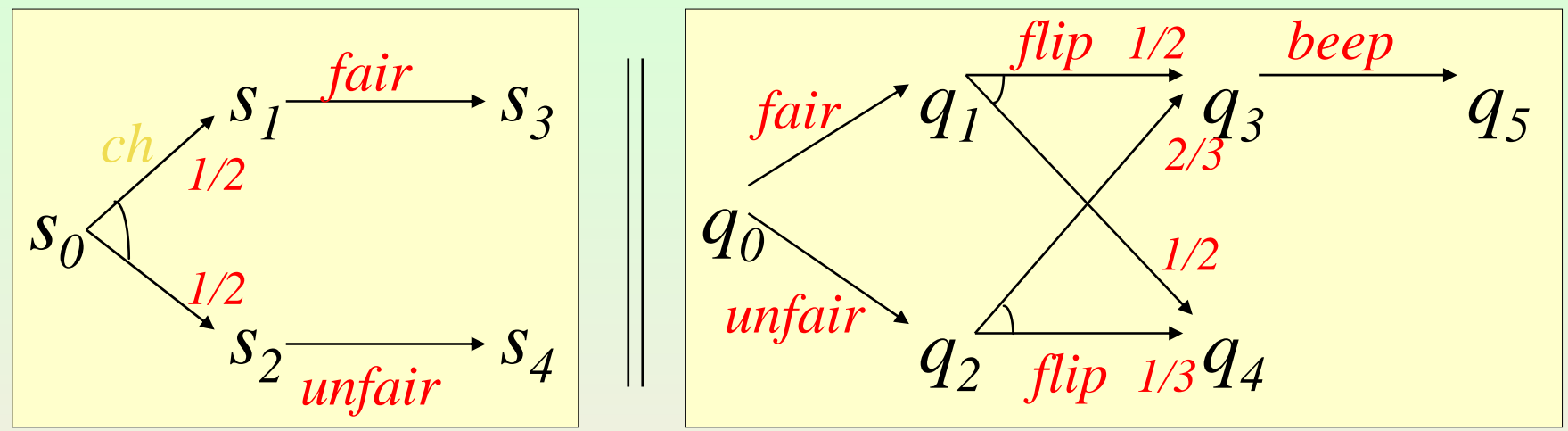
$E = \{n, d, choc, coffee\}$



$E = \{n, d, choc, coffee\}$



# Ex. Composition of Probabilistic Automata



# Projections

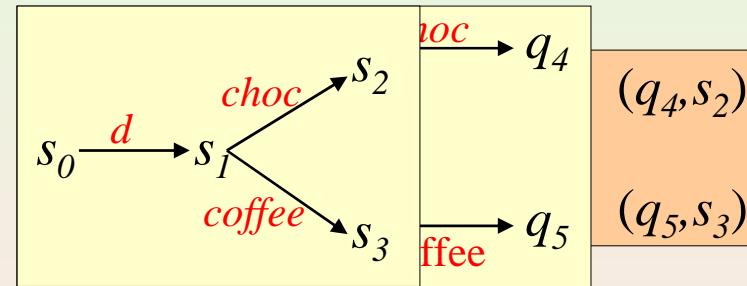
Let  $\alpha$  be an execution of  $A_1 \parallel A_2$

$\alpha = (q_0, s_0) \mathbf{d} (q_2, s_1) \mathbf{ch} (q_3, s_1) \mathbf{coffee} (q_5, s_3)$

What are the contributions of  $A_1$  and  $A_2$ ?

$\pi_1(\alpha) \equiv q_0 \mathbf{d} q_2 \mathbf{ch} q_3 \mathbf{coffee} q_5$

$\pi_2(\alpha) \equiv s_0 \mathbf{d} s_1 \mathbf{coffee} s_3$



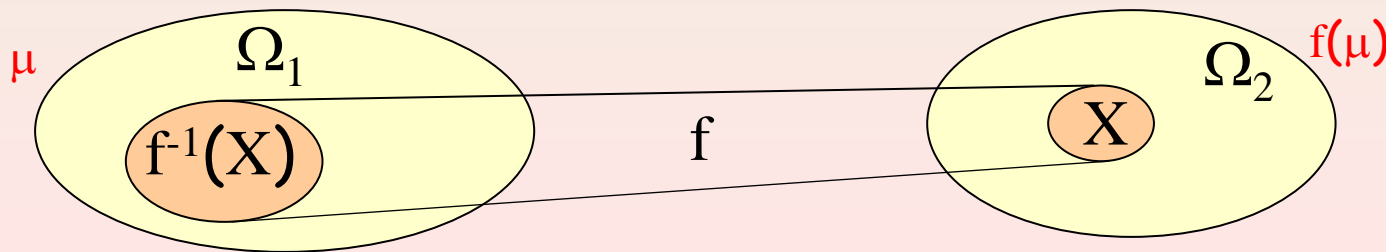
## Theorem

$\alpha \in \text{execs}(A_1 \parallel A_2)$  iff  $\forall i \in \{1,2\} \pi_i(\alpha) \in \text{execs}(A_i)$



# Measure Theory: Image Measure

- Measurable function from  $(\Omega_1, \mathcal{F}_1)$  to  $(\Omega_2, \mathcal{F}_2)$ 
  - Function  $f$  from  $\Omega_1$  to  $\Omega_2$
  - For each element  $X$  of  $\mathcal{F}_2$ ,  $f^{-1}(X) \in \mathcal{F}_1$
- Image measure  $f(\mu)$ 
  - $f(\mu)(X) = \mu(f^{-1}(X))$



# Projections

The projection function is measurable

$\pi(\mu)$  : image measure under  $\pi$  of  $\mu$

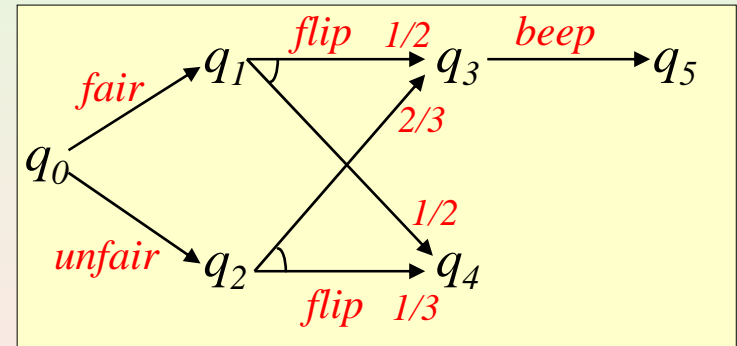
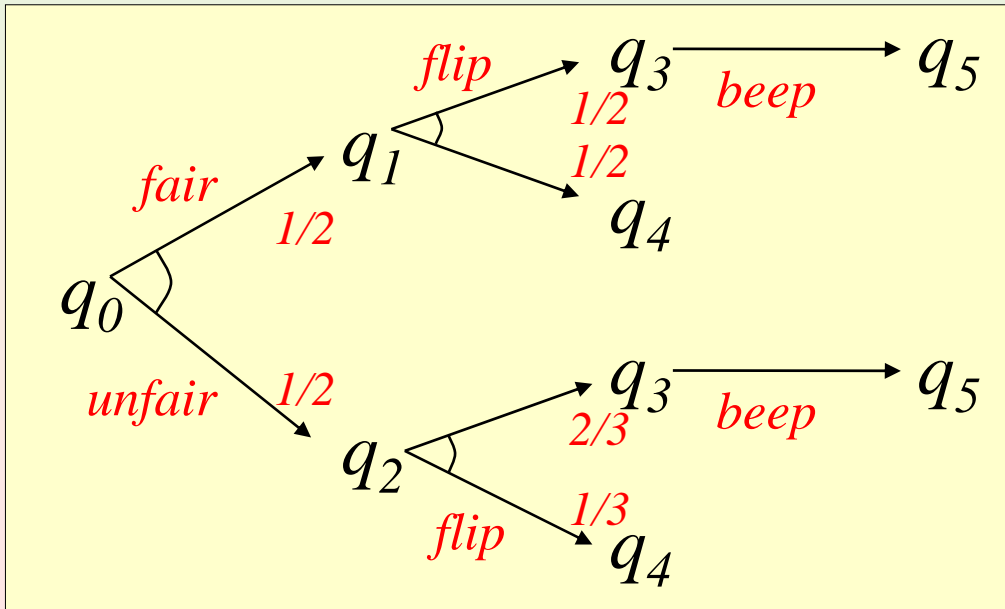
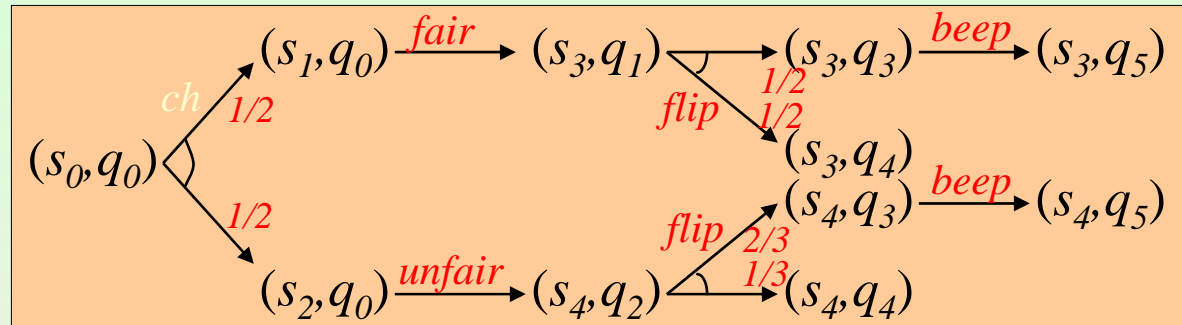
## Theorem

If  $\mu$  is a probabilistic execution of  $A_1 \parallel A_2$   
then

$\pi_i(\mu)$  is a probabilistic execution of  $A_i$

# Example: Projection

Projection onto right component



Note that the scheduler is randomized

# Use of Projections

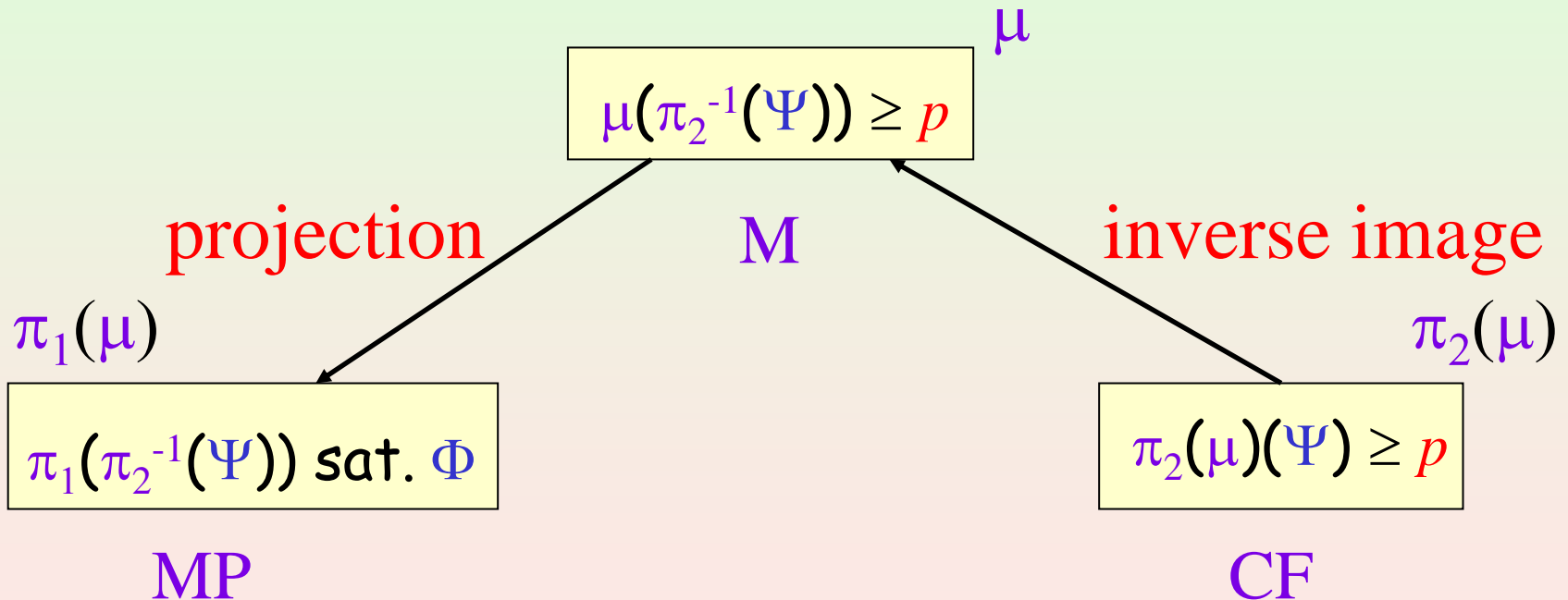
- Let  $M = MP || CF$
- Suppose that  $MP$  satisfies  $\Phi$  provided that the environment ( $CF$ ) satisfies  $\Psi$
- Suppose that  $CF$  satisfies  $\Psi$  with probability  $p$
- Can I conclude that  $M$  satisfies  $\Phi$  with probability  $p$ ?

$$\frac{MP \models \Psi \Rightarrow \Phi \quad CF \models [\Psi]_{\geq p}}{M \models [\Phi]_{\geq p}}$$

- This example is taken from a real case study [PLS01]
  - Randomized consensus protocol of Aspnes and Herlihy [AH90]
  - $MP$  is a complex non randomized protocol
  - $CF$  is a relatively simple randomized coin flipper

# Formal Argument

Let  $\mu$  be a probabilistic execution of  $M$ .



# Bisimulation Relations

---

We have the following objectives

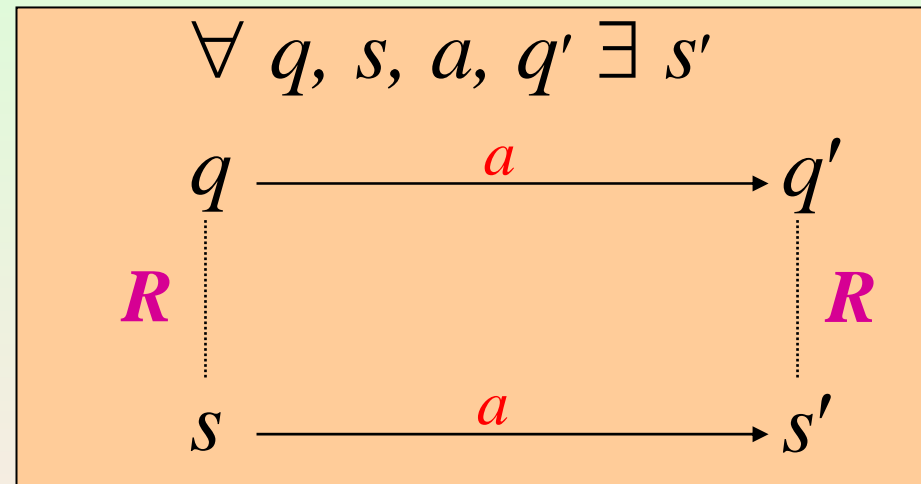
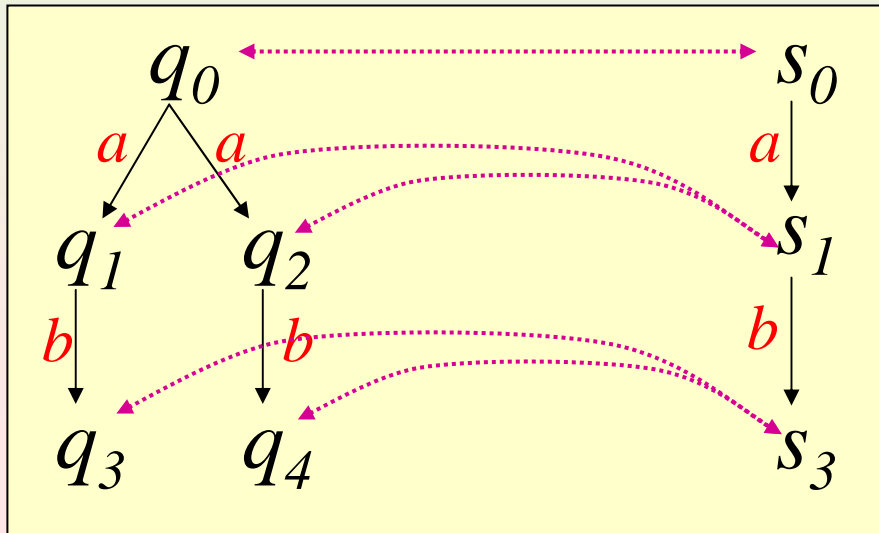
- Same definitional style as for automata
  - Where are the key differences?
  - Keep definitions simple
- Uniform treatment
  - The literature is not uniform
  - This causes a lot of confusion
  - How can we see everything from a single point of view?



# Strong Bisimulation on Automata

Strong bisimulation between  $A_1$  and  $A_2$

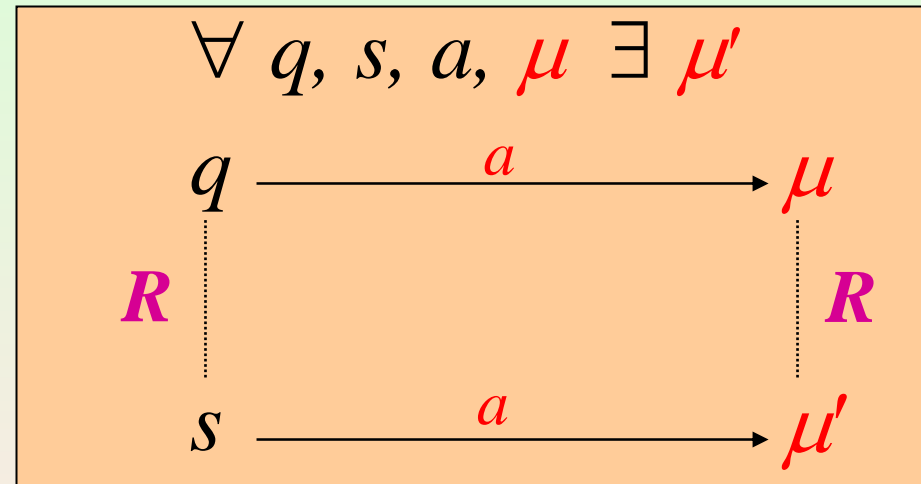
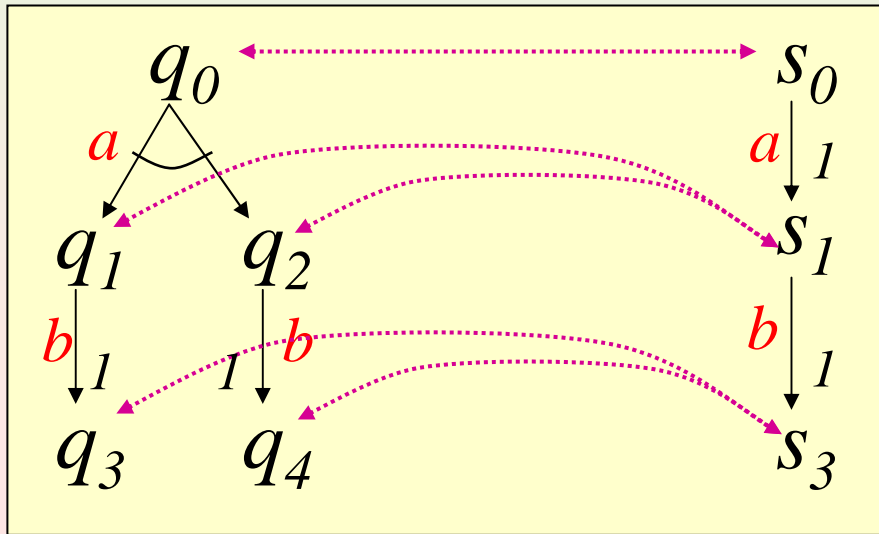
Relation  $R \subseteq Q \times Q$ ,  
 $Q = Q_1 \uplus Q_2$ , such that



# Strong Bisimulation on Probabilistic Automata

Strong bisimulation between  $A_1$  and  $A_2$

Relation  $R \subseteq Q \times Q$ ,  
 $Q = Q_1 \uplus Q_2$ , such that



$$\mu R \mu' \quad [\text{LS89}]$$

$$\Leftrightarrow$$

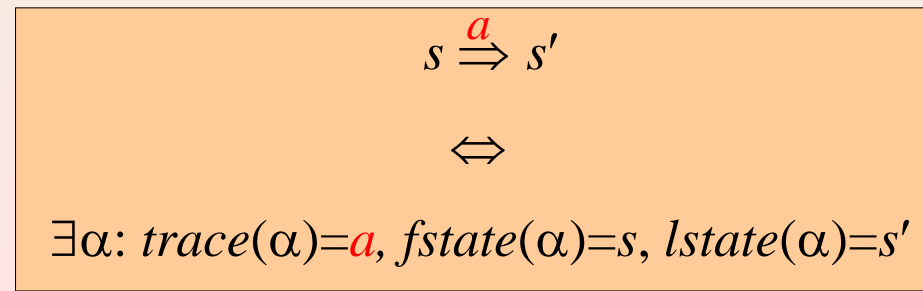
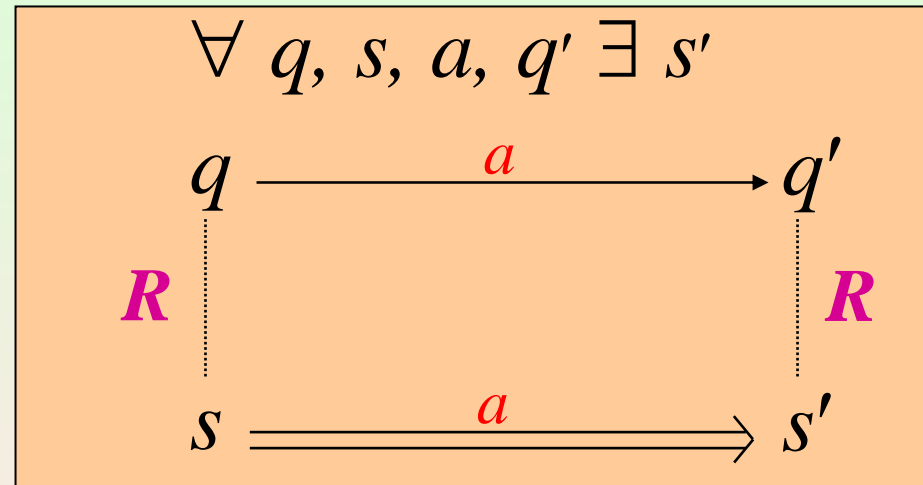
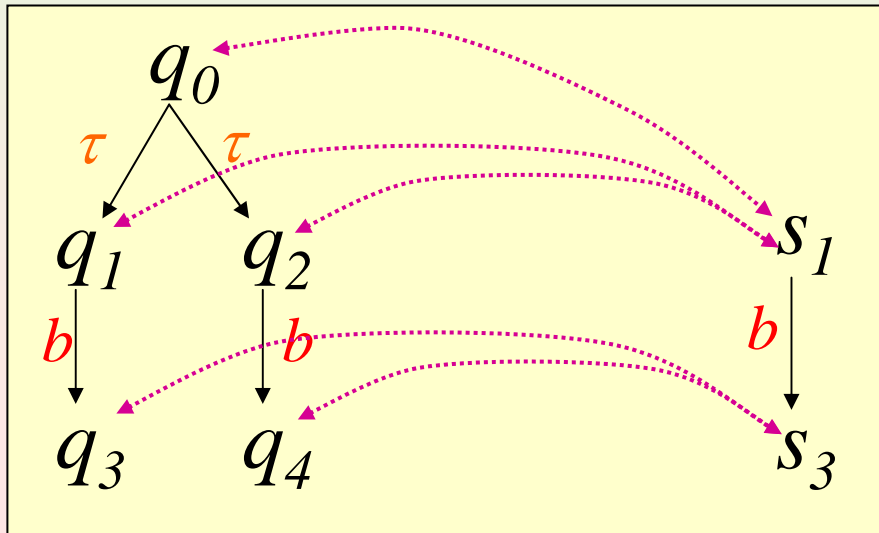
$$\forall C \in Q/R . \mu(C) = \mu'(C)$$



# Weak Bisimulation on Automata

Weak bisimulation between  $A_1$  and  $A_2$

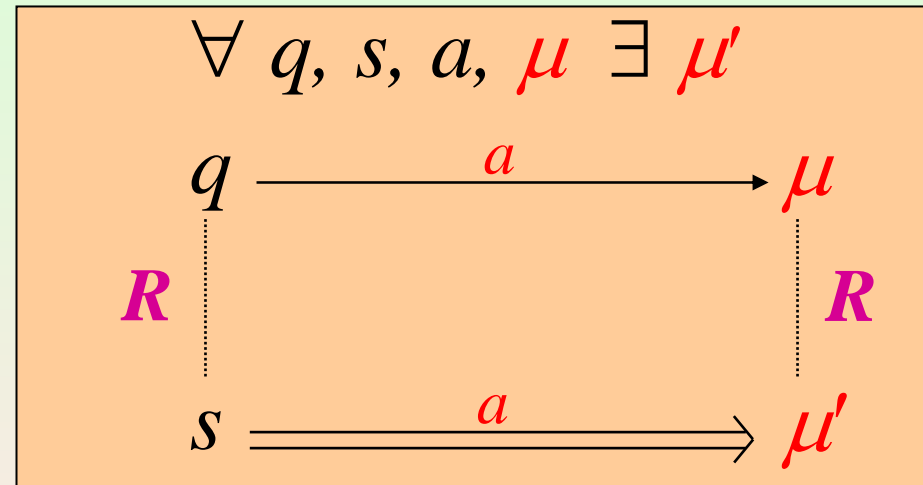
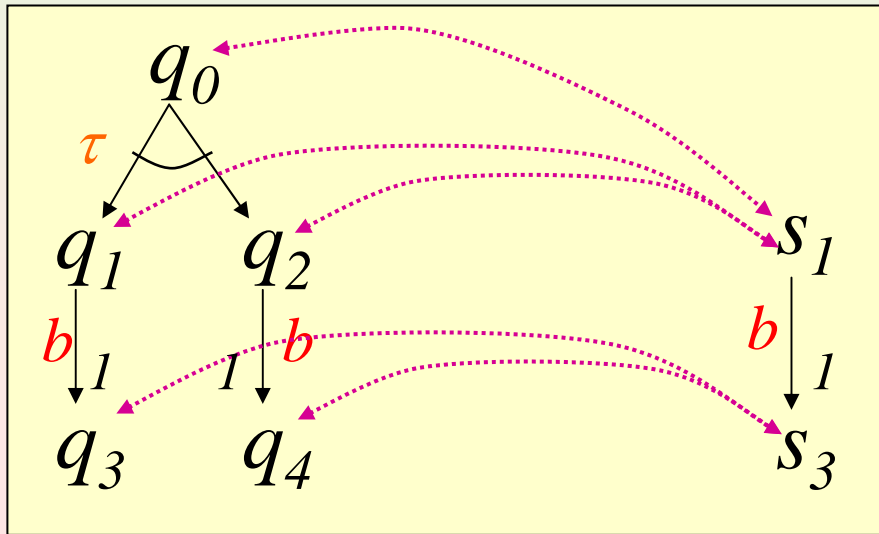
Relation  $R \subseteq Q \times Q$ ,  
 $Q = Q_1 \uplus Q_2$ , such that



# Weak bisimulation on Probabilistic Automata

Weak bisimulation between  $A_1$  and  $A_2$

Relation  $R \subseteq Q \times Q$ ,  
 $Q = Q_1 \uplus Q_2$ , such that



$$\mu R \mu' \quad [\text{LS89}]$$

$$\Leftrightarrow$$

$$\forall C \in Q/R . \mu(C) = \mu'(C)$$

# Weak Transition

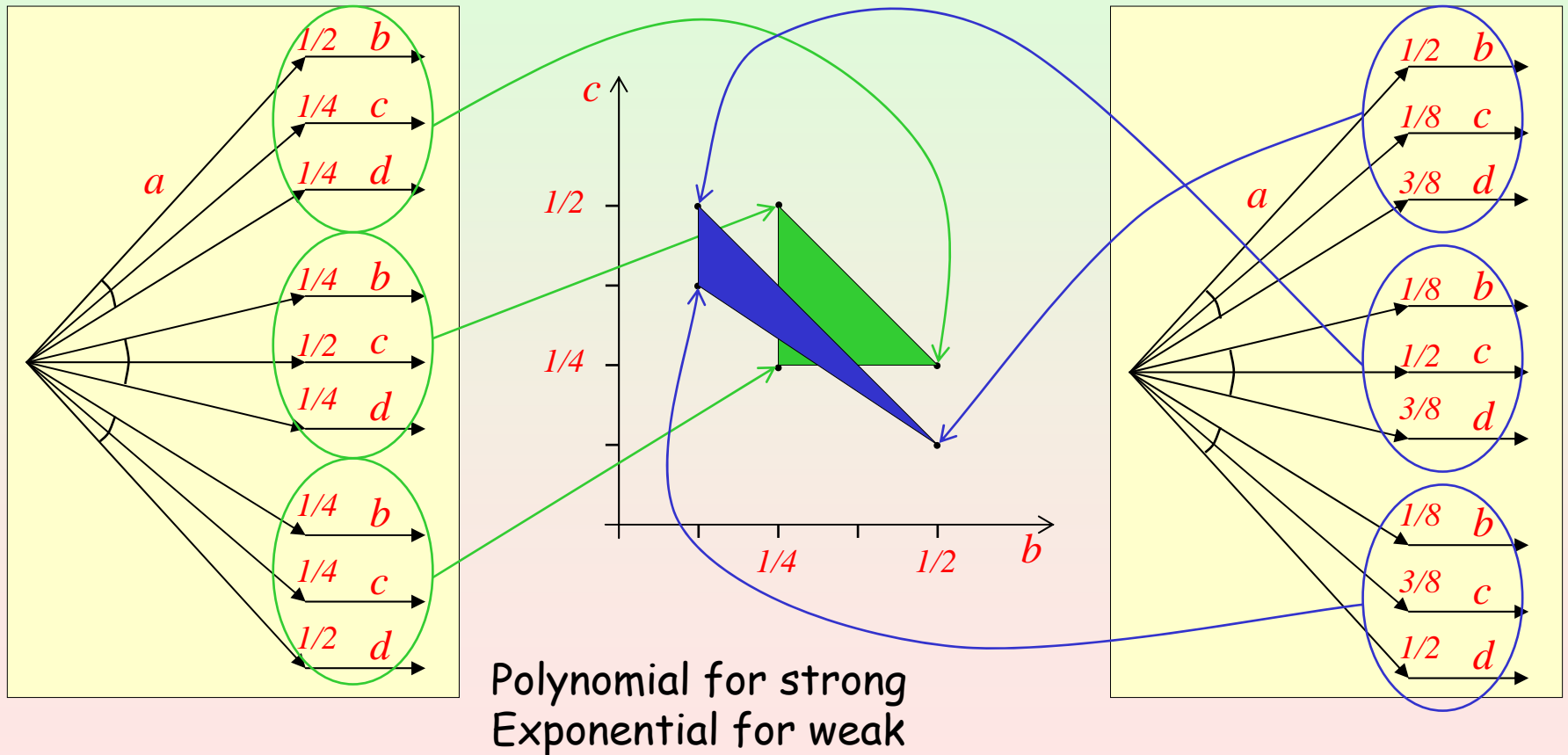
$$q \xRightarrow{a} \rho$$

There is a probabilistic execution  $\mu$  such that

- $\mu(exec^*) = 1$  (it is finite)
- $trace(\mu) = \delta(a)$  (its trace is  $a$ )
- $fstate(\mu) = \delta(q)$  (it starts from  $q$ )
- $lstate(\mu) = \rho$  (it leads to  $\rho$ )

$$q \xRightarrow{a} s \text{ iff } \exists \alpha: trace(\alpha)=a, fstate(\alpha)=q, lstate(\alpha)=s$$

# Decision Procedures



# Trace Distributions

The *trace* function is measurable

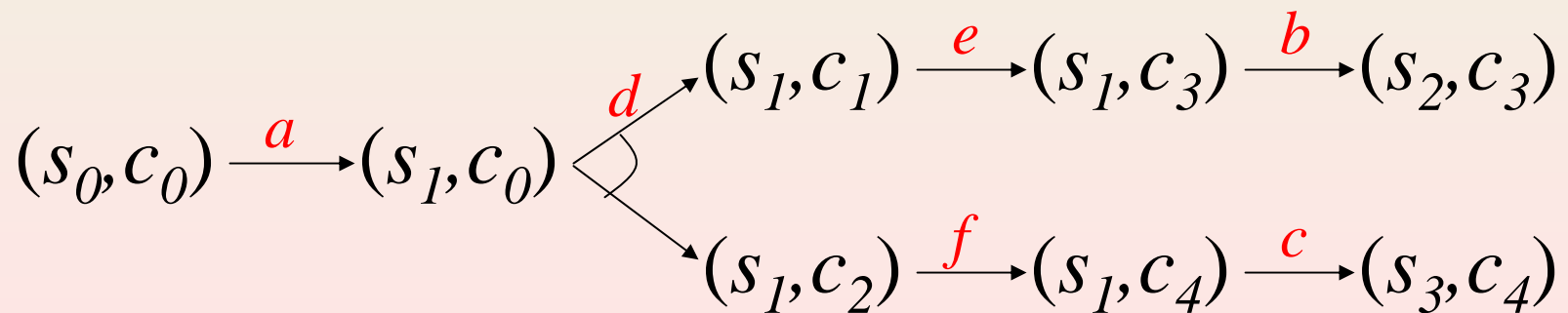
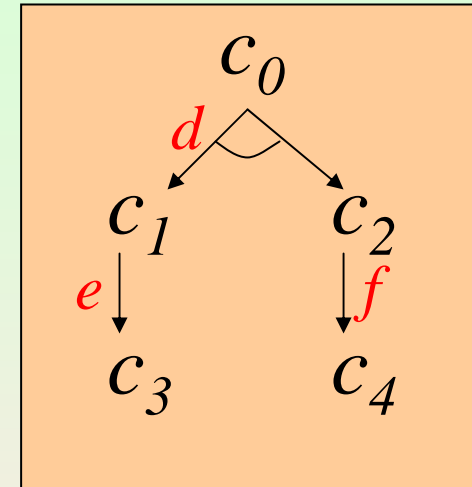
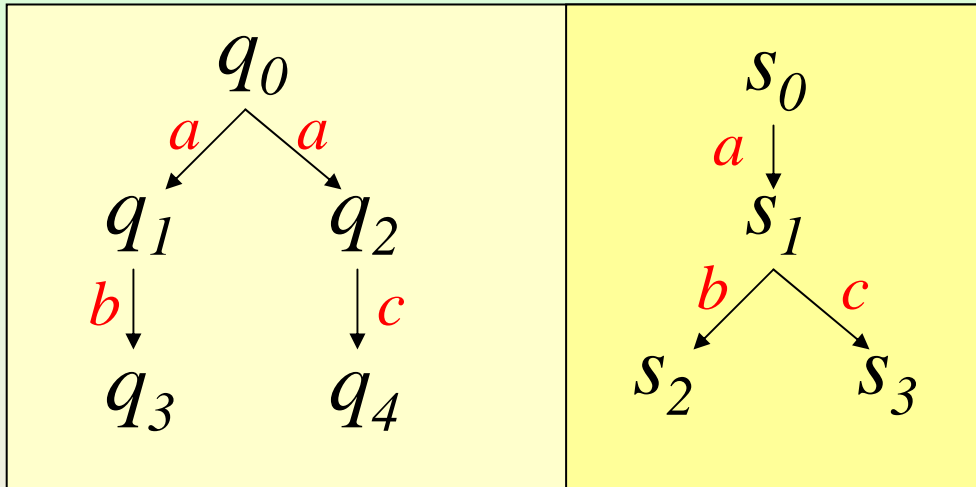
Trace distribution of  $\mu$

$tdist(\mu)$  : image measure under *trace* of  $\mu$

Trace distribution inclusion preorder

$A_1 \leq_{TD} A_2$  iff  $tdists(A_1) \subseteq tdists(A_2)$

# Trace Distribution Inclusion is not Compositional



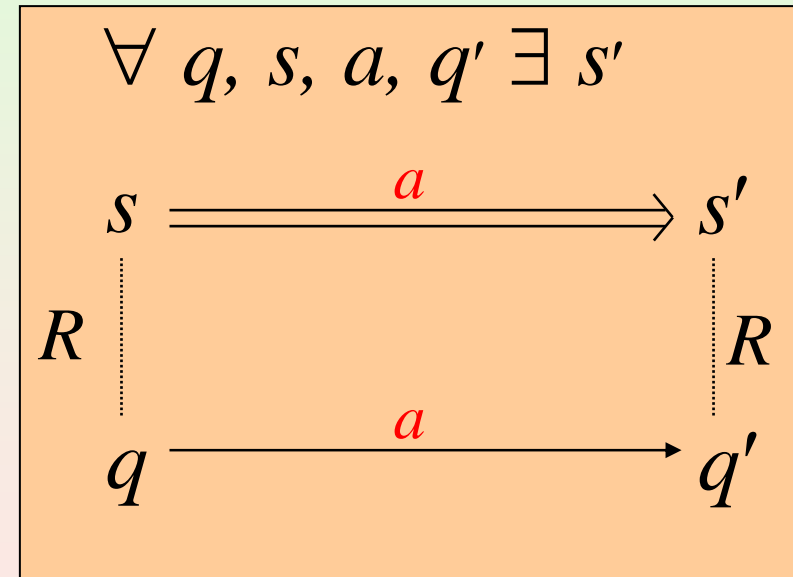
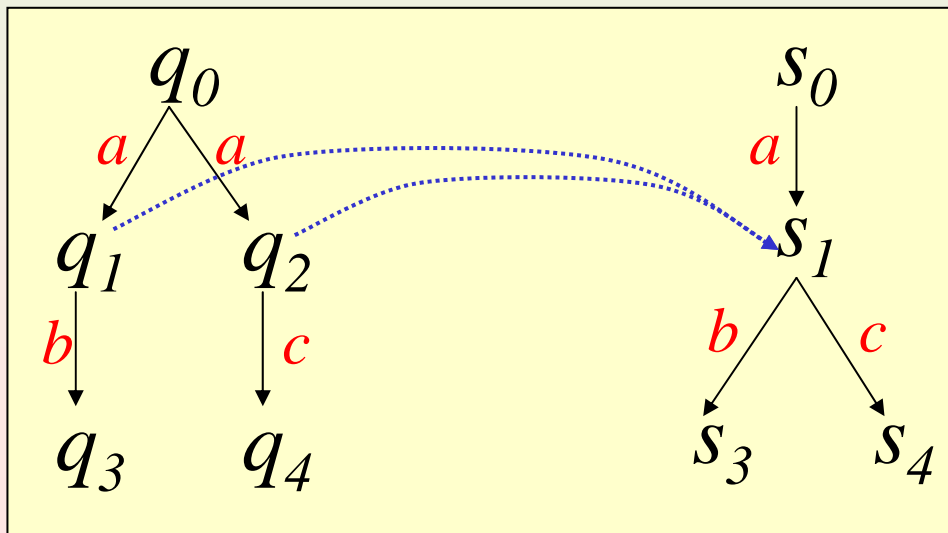
# How to Get Compositionality

- Restrict the power of composition
  - Probabilistic reactive modules [AHJ01]
  - Switched probabilistic I/O automata [CLSV04]
- Trace Distribution Precongruence
  - Coarsest precongruence included in preorder
  - Alternative characterizations
    - Principal context [Seg95]
    - Testing [Seg96]
    - Forward simulations [LSV03]



# Forward Simulations (Automata)

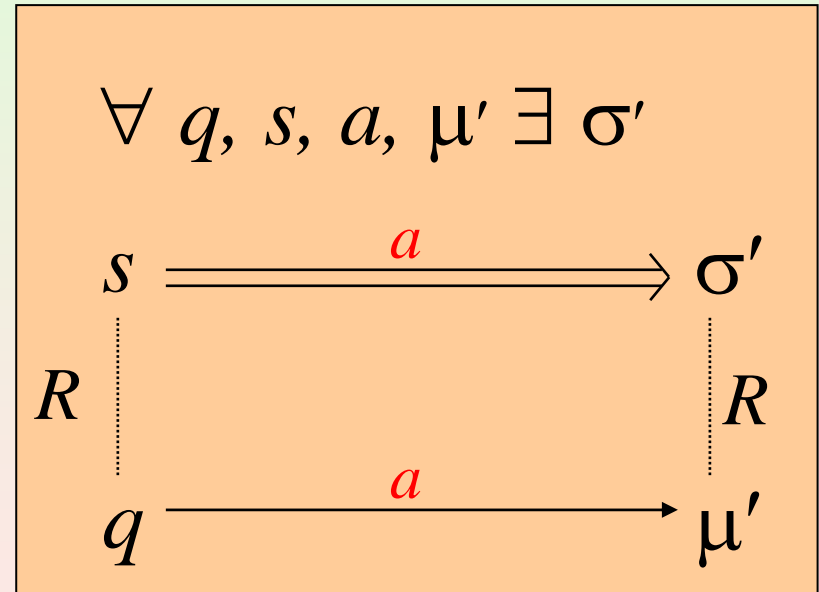
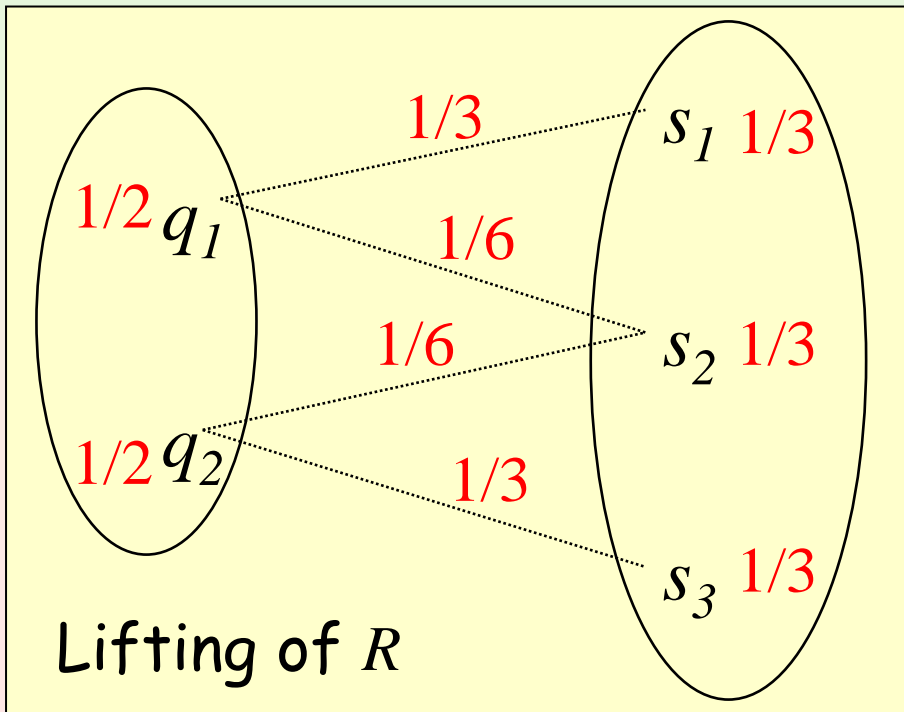
Forward simulation from  $A_1$  to  $A_2$  ( $A_1 \leq_F A_2$ )  
Relation  $R \subseteq Q_1 \times Q_2$  such that



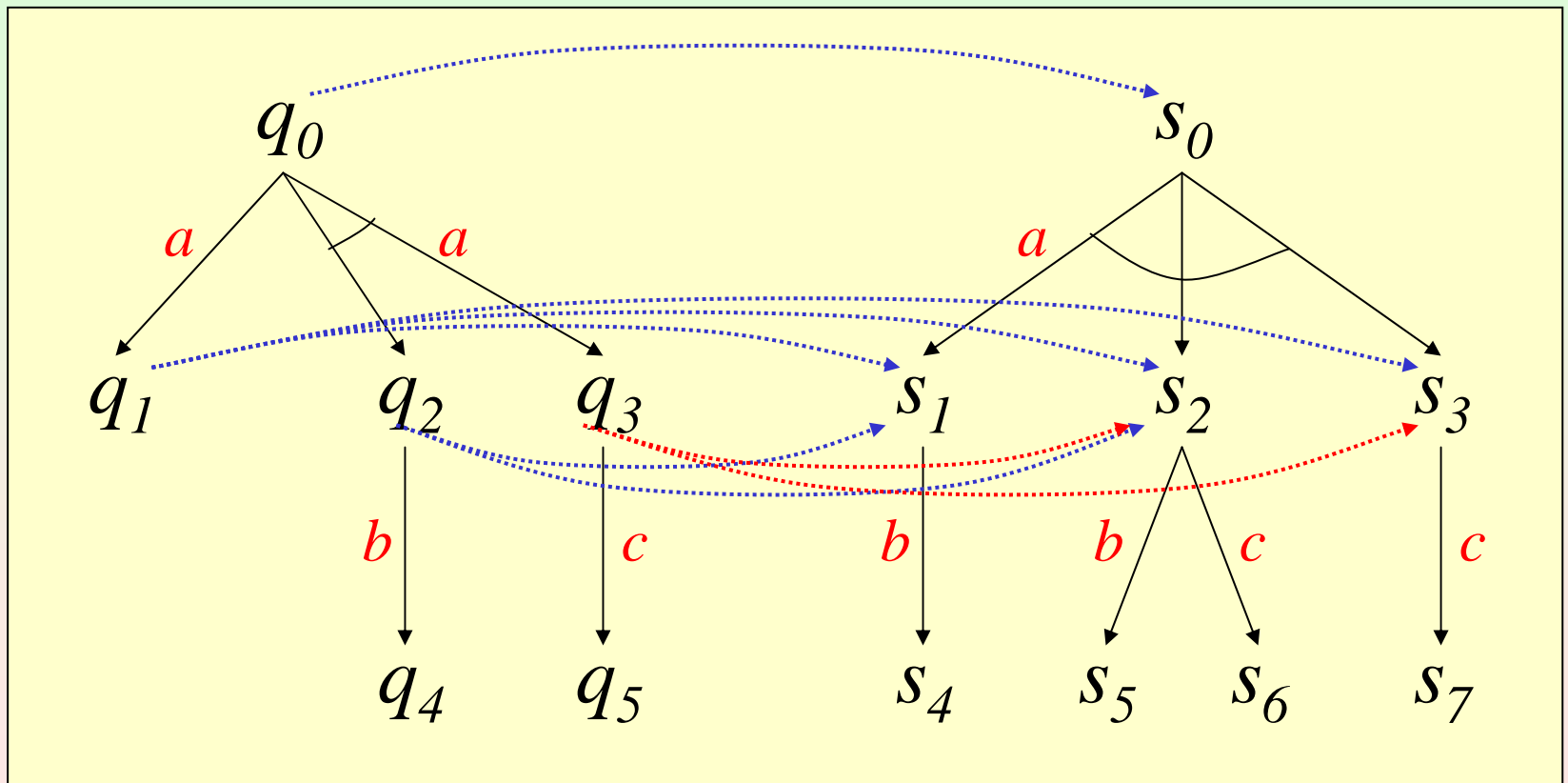


# Forward Simulations

Forward simulation from  $A_1$  to  $A_2$  ( $A_1 \leq_F A_2$ )  
 Relation  $R \subseteq Q_1 \times Q_2$  such that

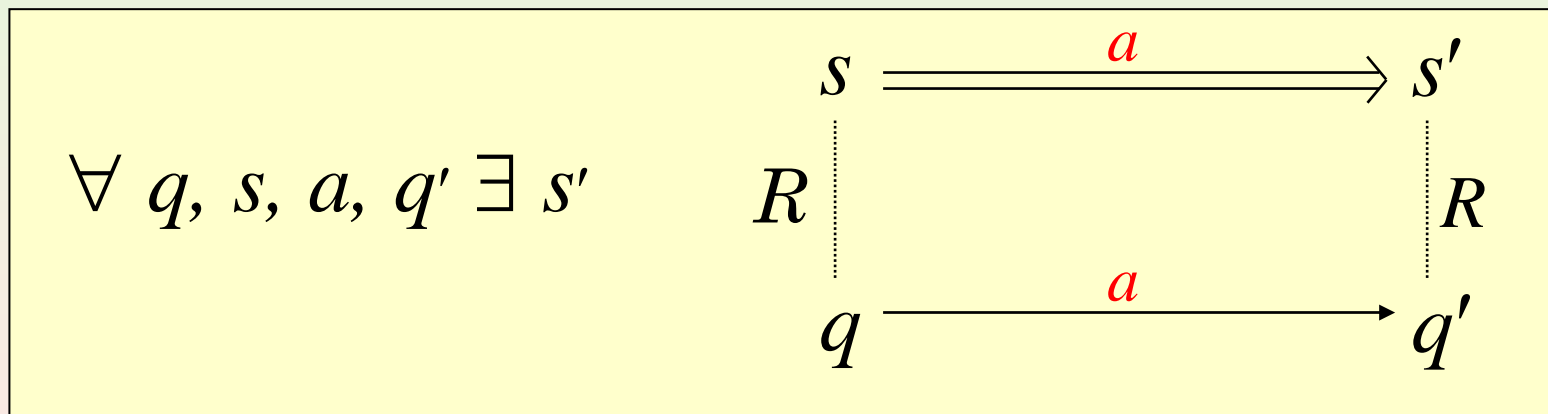


# Example: Simulations



# Characterization: Forward Simulations (Automata)

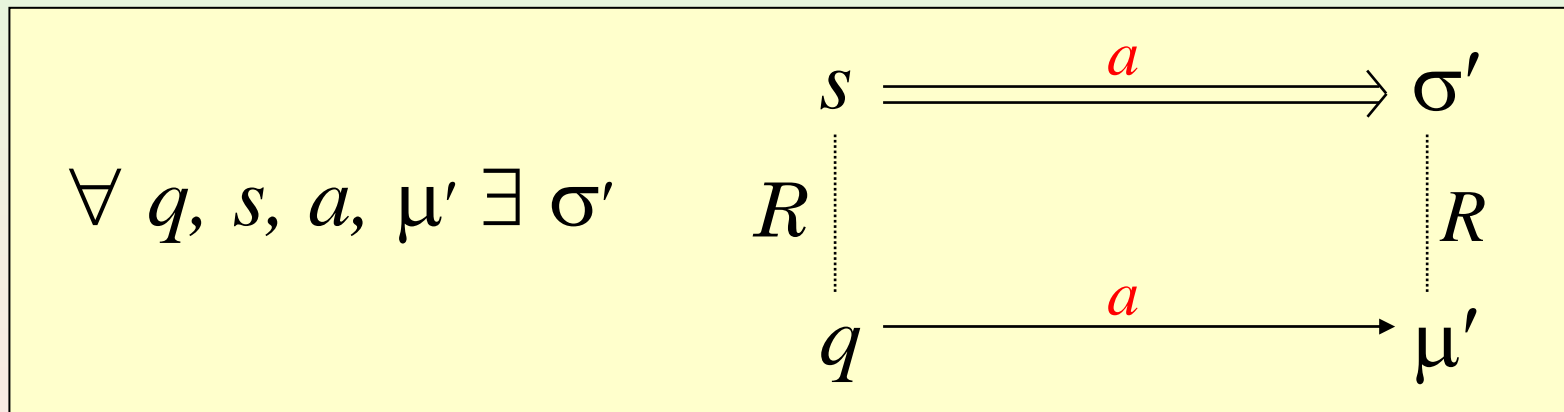
Forward simulation from  $A_1$  to  $A_2$  ( $A_1 \leq_F A_2$ )  
Relation  $R \subseteq Q_1 \times Q_2$  such that



**Theorem** [LSV02]  $A_1 \leq_F A_2$  iff  $A_1 \leq_{\text{TDC}} A_2$

# Characterization: Forward Simulations

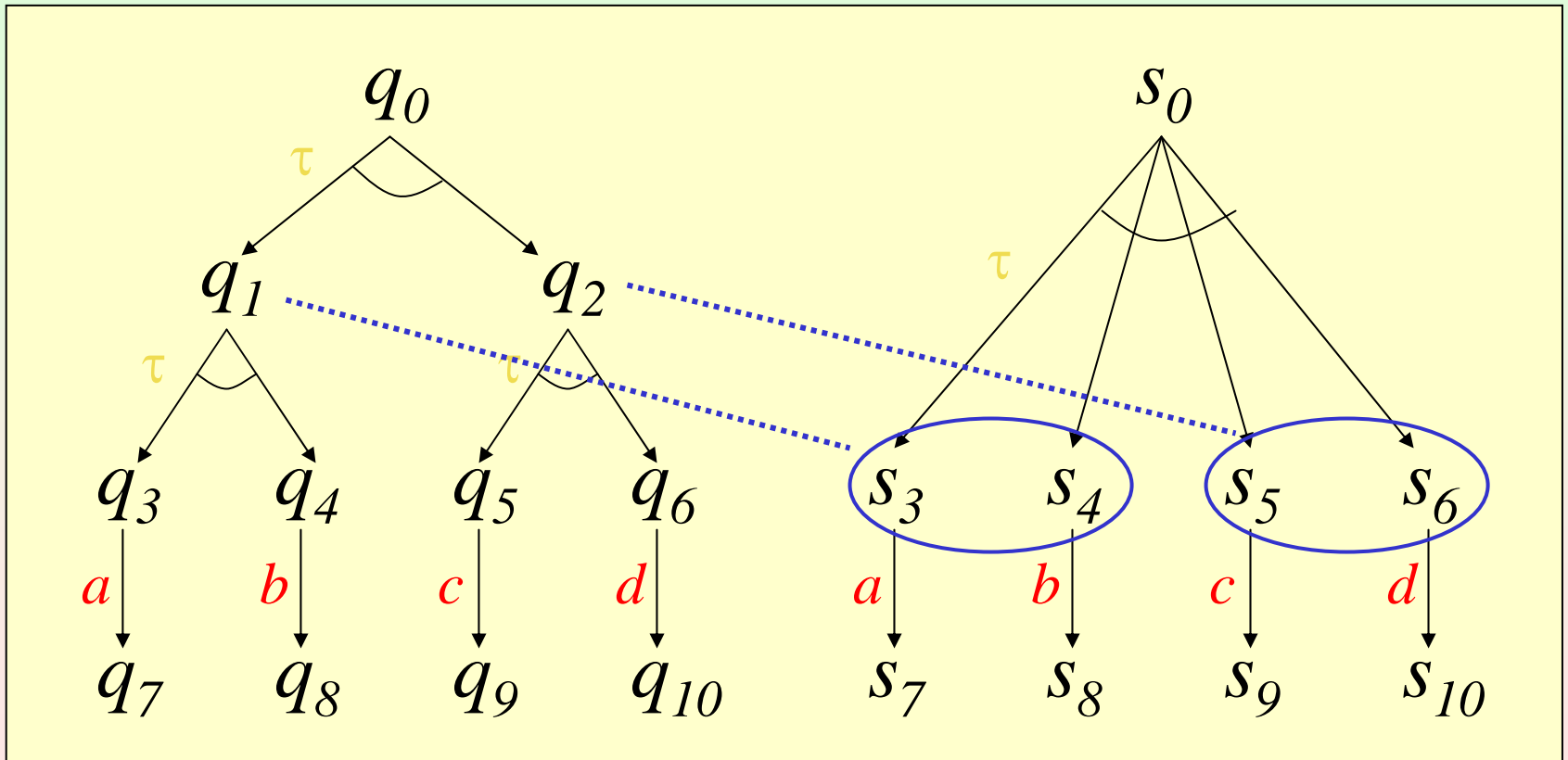
Forward simulation from  $A_1$  to  $A_2$  ( $A_1 \leq_F A_2$ )  
Relation  $R \subseteq Q_1 \times Q_2$  such that



**Theorem**  $A_1 \leq_F A_2$  implies  $A_1 \leq_{\text{TDC}} A_2$

# Example:

## Failure of Weak Forward Simulations



# Characterization: Probabilistic Forward Simulations

Forward simulation from  $A_1$  to  $A_2$  ( $A_1 \leq_{\text{PF}} A_2$ )  
Relation  $R \subseteq Q_1 \times \text{Disc}(Q_2)$  such that

$$\forall q, \sigma, a, \mu' \exists \sigma'', \sigma' \quad \begin{array}{ccc} \sigma & \xrightarrow{a} & \sigma'' \equiv \sigma' \\ \vdots R & & \vdots R \\ q & \xrightarrow{a} & \mu' \end{array}$$

**Theorem** [LSV02]  $A_1 \leq_{\text{PF}} A_2$  iff  $A_1 \leq_{\text{TDC}} A_2$

# Other Important Topics

---

- Axiomatizations
- Logical characterizations
- Probabilistic language inclusion
- Simulation relations
- Metrics
- Testing
- Model checking
- Timed models
  - Timed Probabilistic Automata
  - Stochastic Transition Systems