### **Labelled Markov Processes**

Prakash Panangaden

School of Computer Science McGill University



 Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is..



- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is..
- Wait!! What are Labelled Transition Systems?



Systems that can be in a state and have transitions between states; these transitions are triggered by actions. The transitions are annotated by labels that name the action. [Invented by R. Keller 1976]



- Systems that can be in a state and have transitions between states; these transitions are triggered by actions. The transitions are annotated by labels that name the action. [Invented by R. Keller 1976]
- The transitions can be nondeterministic.



>

- Systems that can be in a state and have transitions between states; these transitions are triggered by actions. The transitions are annotated by labels that name the action. [Invented by R. Keller 1976]
- The transitions can be nondeterministic.
- Intended to model communication and concurrency; the notion of observation is very different from what one uses in automata theory.



<

- Systems that can be in a state and have transitions between states; these transitions are triggered by actions. The transitions are annotated by labels that name the action. [Invented by R. Keller 1976]
- The transitions can be nondeterministic.
- Intended to model communication and concurrency; the notion of observation is very different from what one uses in automata theory.
- We do not see the states, we see the actions and we observe when actions are rejected by the system.

<

>

+

 Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.



- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* no probabilities associated with environment behaviour.



- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* no probabilities associated with environment behaviour.
- We observe the interactions not the internal states.

<

- Labelled Markov processes are probabilistic versions of labelled transition systems. Labelled transition systems where the final state is governed by a probability distribution - no other indeterminacy.
- All probabilistic data is *internal* no probabilities associated with environment behaviour.
- We observe the interactions not the internal states.
- In general, the state space of a labelled Markov process may be a *continuum*.

+

<

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

probabilistic process algebra with recursion,





Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.



Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation;
  e.g. cell phones

+

>

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation;
  e.g. cell phones

+

>

<

performance modelling,

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation;
  e.g. cell phones

+

>

- performance modelling,
- population growth models, stock market

Model and reason about systems with *continuous* state spaces or continuous time evolution or both.

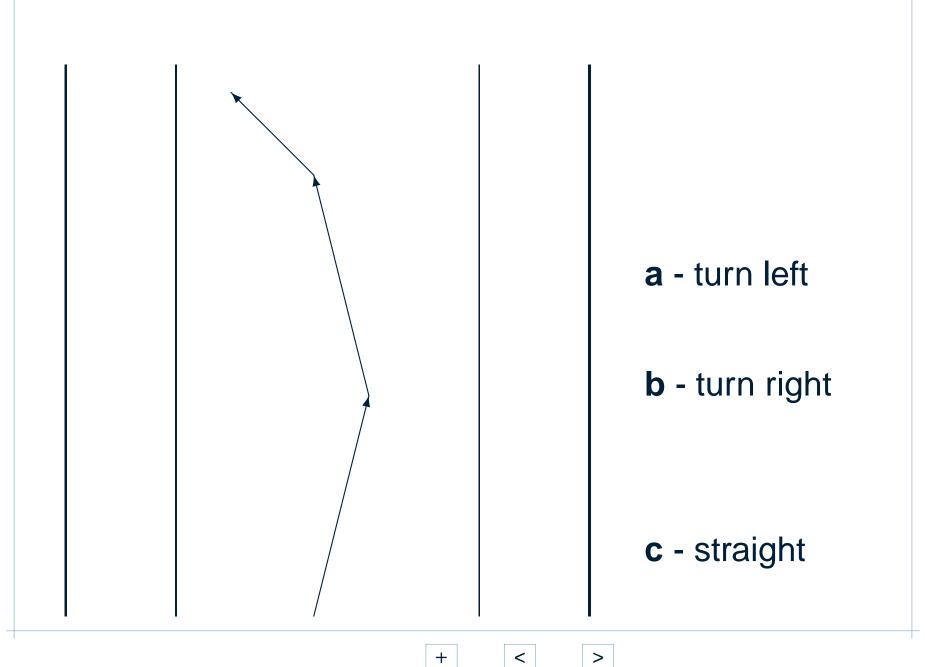
- probabilistic process algebra with recursion,
- hybrid control systems; e.g. flight management systems.
- telecommunication systems with spatial variation;
  e.g. cell phones

+

>

- performance modelling,
- population growth models, stock market
- continuous time systems

## **An Example**



+

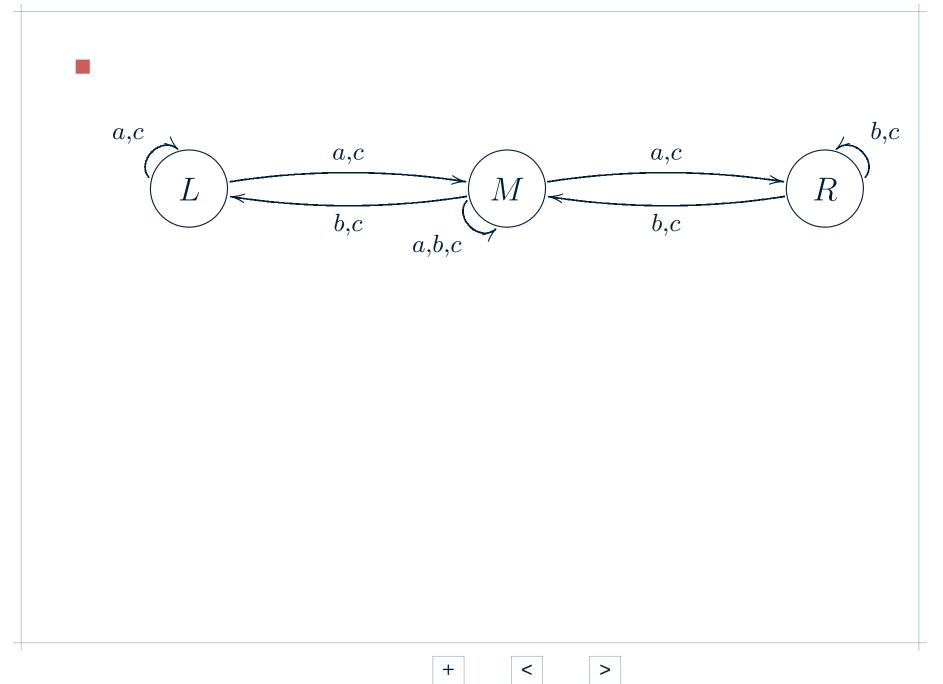
# Actions

a - turn left, b - turn right, c - keep on course The actions move the craft sideways with some probability distributions on how far it moves. The craft may "drift" even with c. The action a (b) must be disabled when the craft is too near the left (right) boundary.

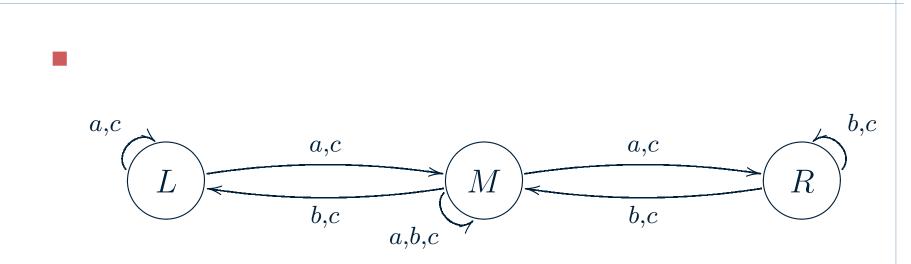
<



#### **Schematic of Example**



### **Schematic of Example**



This picture is misleading: unless very special conditions hold the process cannot be compressed into an *equivalent* (?) finite-state model. In general, the transition probabilities should depend on the position.

+

<

- This is a toy model but exemplifies the issues.
- Can be used for reasoning much better if we could have a finite-state version.
- Why not discretize right away and never worry about the continuous case? Because we lose the ability to refine the model later.
- A better model would be to base it on rewards and think about finiding optimal policies as in AI literature.

### **Probabilistic Transition Systems**

 A transition system with probabilities and actions (labels) associated with the transitions.



## **Probabilistic Transition Systems**

 A transition system with probabilities and actions (labels) associated with the transitions.

$$(S, \mathsf{L}, \forall a \in \mathsf{L} \ T_a : S \times S \longrightarrow [0, 1])$$



# **Probabilistic Transition Systems**

 A transition system with probabilities and actions (labels) associated with the transitions.

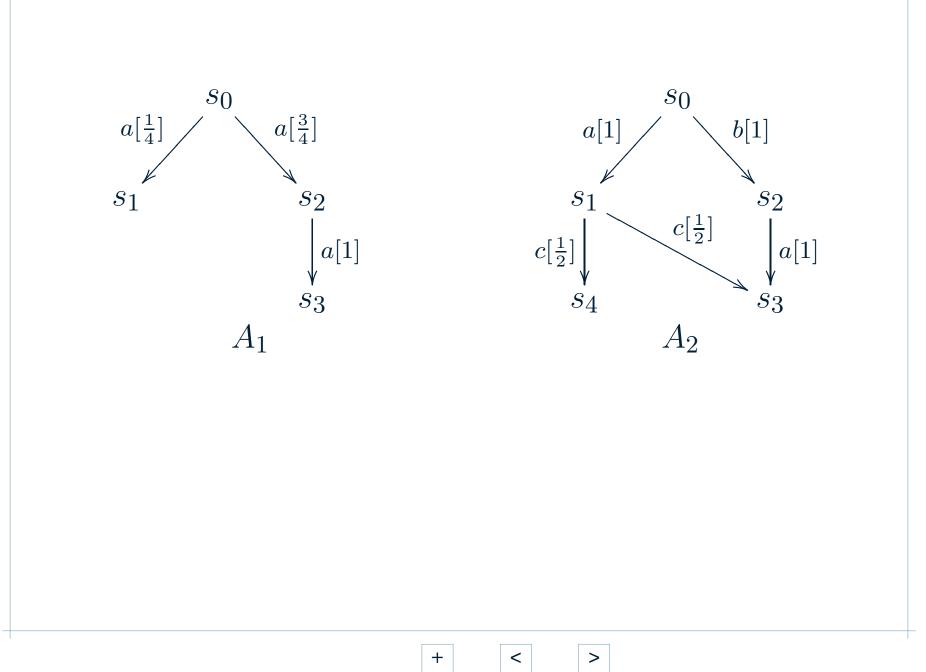
$$(S, \mathsf{L}, \forall a \in \mathsf{L} \ T_a : S \times S \longrightarrow [0, 1])$$

 The model is *reactive*: All probabilistic data is *internal* no probabilities associated with environment behaviour.



<

#### **Examples of PTSs**



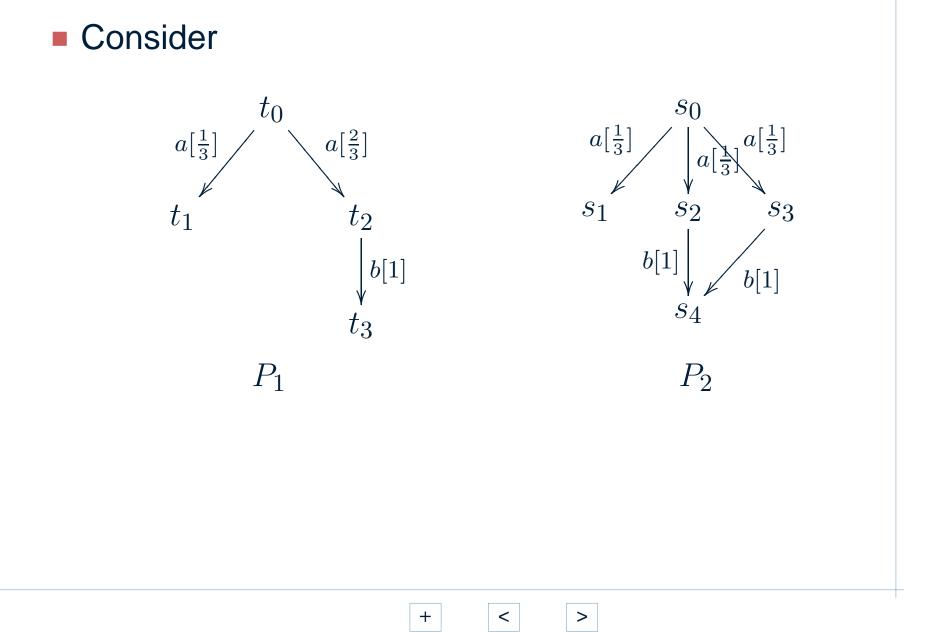
If two states behave in exactly the same way they can be combined.

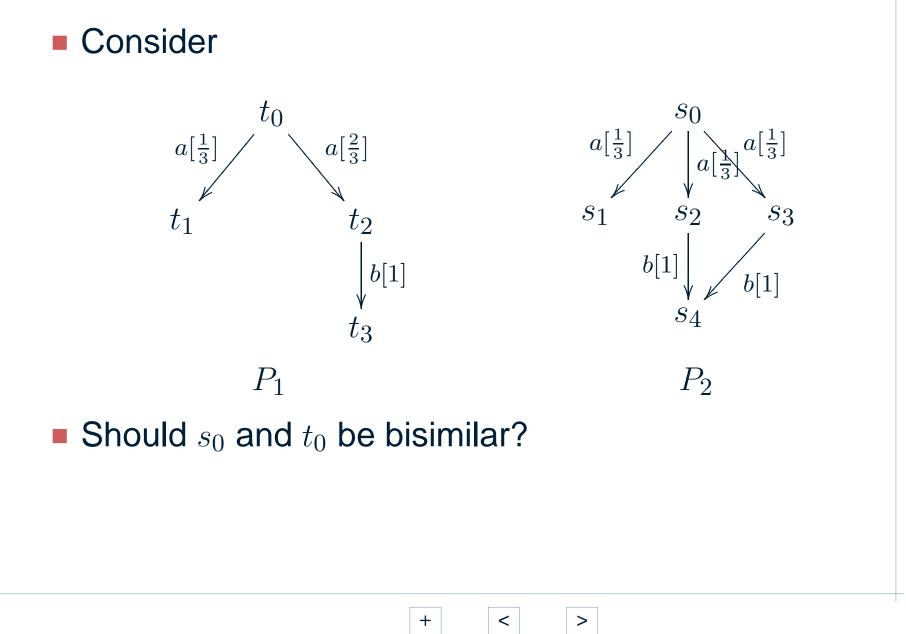


- If two states behave in exactly the same way they can be combined.
- In queueing theory there was a notion of lumpability of Markov chains (with no labels).

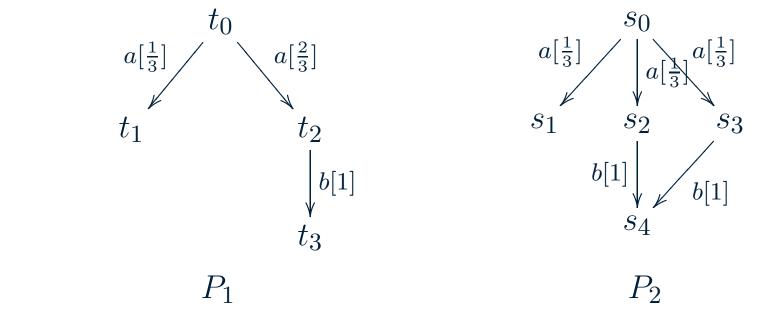


- If two states behave in exactly the same way they can be combined.
- In queueing theory there was a notion of lumpability of Markov chains (with no labels).
- In process algebra (with no probabilities) Park and Milner formulated a notion called bisimulation which captures a very fine notion of process equivalence.









>

• Should  $s_0$  and  $t_0$  be bisimilar?

Yes, but we need to add the probabilities.

+

### **The Official Definition\***

• Let  $S = (S, L, T_a)$  be a PTS. An equivalence relation R on S is a **bisimulation** if whenever sRs', with  $s, s' \in S$ , we have that for all  $a \in A$  and every R-equivalence class,  $A, T_a(s, A) = T_a(s', A)$ .



# **The Official Definition\***

- Let  $S = (S, L, T_a)$  be a PTS. An equivalence relation R on S is a **bisimulation** if whenever sRs', with  $s, s' \in S$ , we have that for all  $a \in A$  and every R-equivalence class,  $A, T_a(s, A) = T_a(s', A)$ .
- The notation  $T_a(s, A)$  means "the probability of starting from s and jumping to a state in the set A."

+

# **The Official Definition\***

- Let  $S = (S, L, T_a)$  be a PTS. An equivalence relation R on S is a **bisimulation** if whenever sRs', with  $s, s' \in S$ , we have that for all  $a \in A$  and every R-equivalence class,  $A, T_a(s, A) = T_a(s', A)$ .
- The notation  $T_a(s, A)$  means "the probability of starting from s and jumping to a state in the set A."
- Two states are bisimilar if there is some bisimulation relation R relating them.

<

Basic fact: There are subsets of R for which no sensible notion of size can be defined.



#### **The Need for Measure Theory\***

- Basic fact: There are subsets of R for which no sensible notion of size can be defined.
- More precisely, there is no translation-invariant measure defined on all the subsets of the reals.



• A discrete-time Markov chain is a finite set S (the state space) together with a transition probability function  $T: S \times S \longrightarrow [0, 1]$ .



- A *discrete-time* Markov chain is a finite set *S* (the state space) together with a transition probability function  $T: S \times S \longrightarrow [0, 1]$ .
- A Markov chain is just a probabilistic automaton; if we add labels we get a PTS.



- A discrete-time Markov chain is a finite set S (the state space) together with a transition probability function  $T: S \times S \longrightarrow [0, 1]$ .
- A Markov chain is just a probabilistic automaton; if we add labels we get a PTS.
- The key property is that the transition probability from s to s' only depends on s and s' and not on the past history of how it got there. This is what allows the probabilistic data to be given as a single matrix T.

<

• A stochastic kernel (Markov kernel) is a function  $h: S \times \Sigma \longrightarrow [0, 1]$  with (a)  $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$  a (sub)probability measure and (b)  $h(\cdot, A) : X \longrightarrow [0, 1]$  a measurable function.



• A stochastic kernel (Markov kernel) is a function  $h: S \times \Sigma \longrightarrow [0, 1]$  with (a)  $h(s, \cdot) : \Sigma \longrightarrow [0, 1]$  a (sub)probability measure and (b)  $h(\cdot, A) : X \longrightarrow [0, 1]$  a measurable function.

Though apparantly asymmetric, these are the stochastic analogues of binary relations



<

+

- A stochastic kernel (Markov kernel) is a function  $h: S \times \Sigma \longrightarrow [0,1]$  with (a)  $h(s, \cdot) : \Sigma \longrightarrow [0,1]$  a (sub)probability measure and (b)  $h(\cdot, A) : X \longrightarrow [0,1]$  a measurable function.
- Though apparantly asymmetric, these are the stochastic analogues of binary relations
- and the uncountable generalization of a matrix.



+

<

#### **Formal Definition of LMPs\***

• An LMP is a tuple  $(S, \Sigma, L, \forall \alpha \in L.\tau_{\alpha})$  where  $\tau_{\alpha} : S \times \Sigma$  $\rightarrow [0, 1]$  is a *transition probability* function such that



### **Formal Definition of LMPs\***

- An LMP is a tuple  $(S, \Sigma, L, \forall \alpha \in L.\tau_{\alpha})$  where  $\tau_{\alpha} : S \times \Sigma$  $\rightarrow [0, 1]$  is a *transition probability* function such that
- $\forall s : S.\lambda A : \Sigma.\tau_{\alpha}(s, A)$  is a subprobability measure and
  - $\forall A : \Sigma . \lambda s : S . \tau_{\alpha}(s, A)$  is a measurable function.

<

+



#### **Larsen-Skou Bisimulation\***

• Let  $S = (S, i, \Sigma, \tau)$  be a labelled Markov process. An equivalence relation R on S is a **bisimulation** if whenever sRs', with  $s, s' \in S$ , we have that for all  $a \in A$  and every R-closed measurable set  $A \in \Sigma$ ,  $\tau_a(s, A) = \tau_a(s', A)$ .

Two states are bisimilar if they are related by a bisimulation relation.



>

<

+

### **Larsen-Skou Bisimulation\***

Let S = (S, i, Σ, τ) be a labelled Markov process. An equivalence relation R on S is a bisimulation if whenever sRs', with s, s' ∈ S, we have that for all a ∈ A and every R-closed measurable set A ∈ Σ, τ<sub>a</sub>(s, A) = τ<sub>a</sub>(s', A).
 Two states are bisimilar if they are related by a

bisimulation relation.

Can be extended to bisimulation between two different LMPs.

+

<

### **Larsen-Skou Bisimulation\***

• Let  $S = (S, i, \Sigma, \tau)$  be a labelled Markov process. An equivalence relation R on S is a **bisimulation** if whenever sRs', with  $s, s' \in S$ , we have that for all  $a \in A$  and every R-closed measurable set  $A \in \Sigma$ ,  $\tau_a(s, A) = \tau_a(s', A)$ . Two states are bisimilar if they are related by a

bisimulation relation.

- Can be extended to bisimulation between two different LMPs.
- Essentially the same as the version that we had before with zigzag morphisms but much closer in spirit to the Larsen-Skou version.

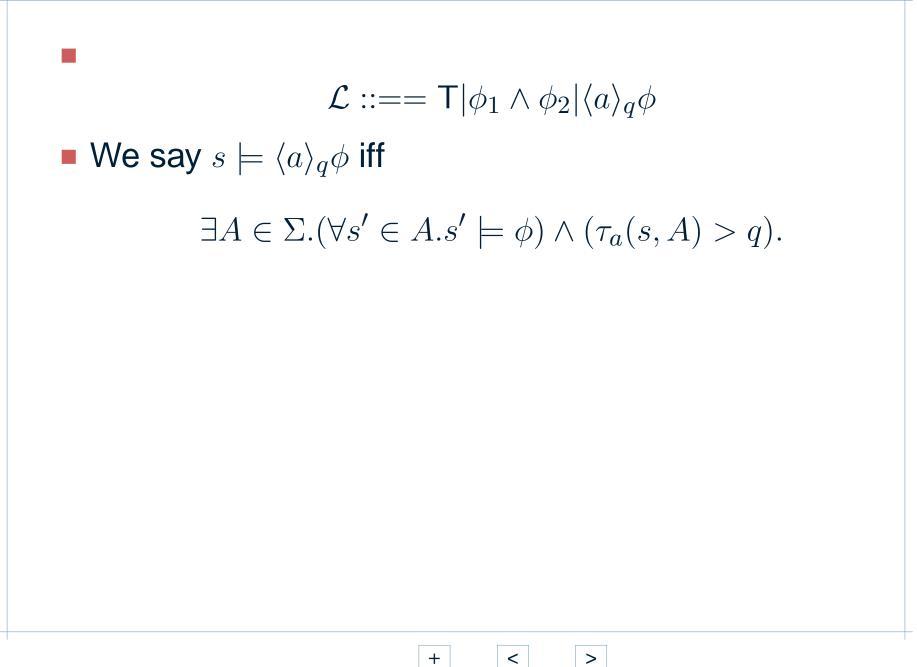


<

#### **Logical Characterization**

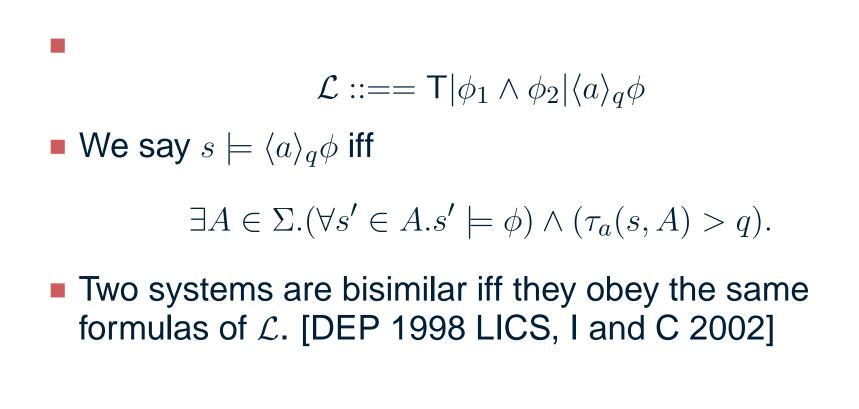


#### **Logical Characterization**



+

#### **Logical Characterization**

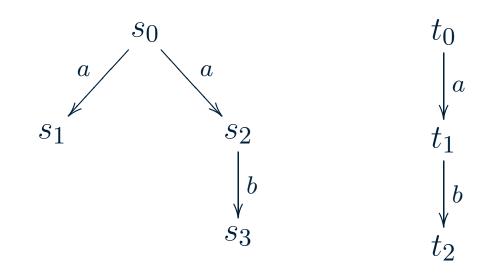




+

<

#### That Cannot be Right?



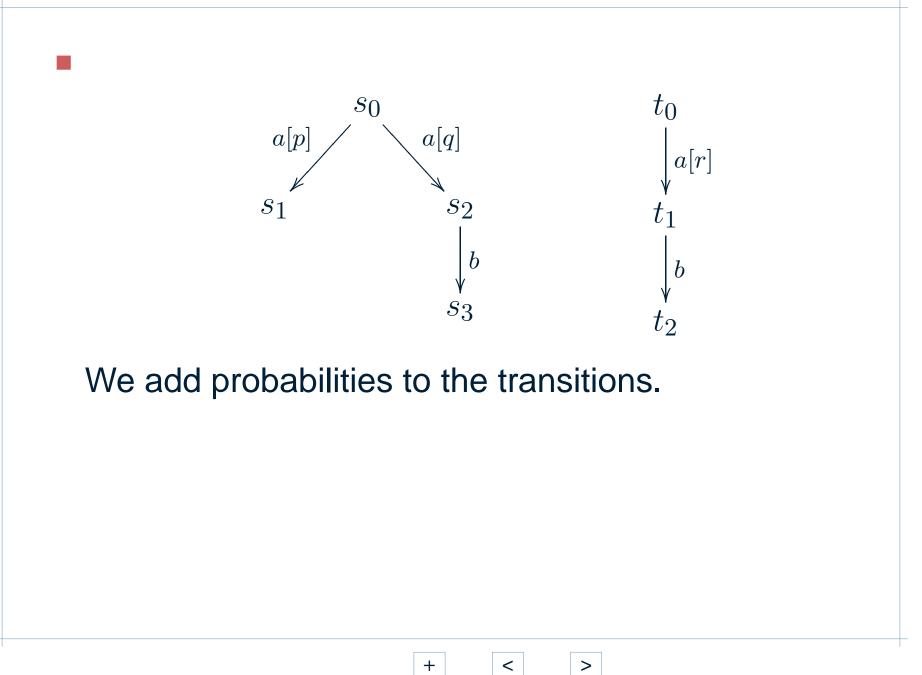
Two processes that cannot be distinguished without negation.

The formula that distinguishes them is  $\langle a \rangle (\neg \langle b \rangle \top)$ .

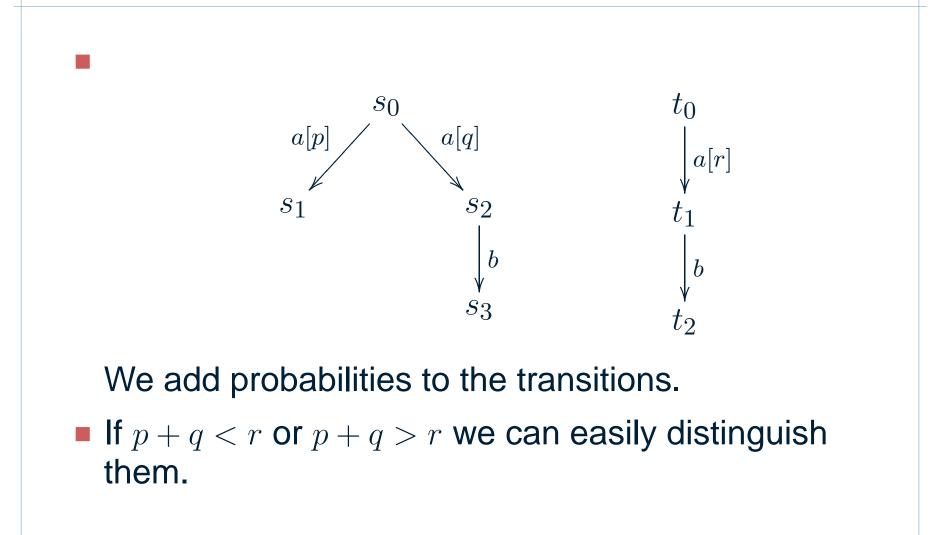
+

>

## But it is!



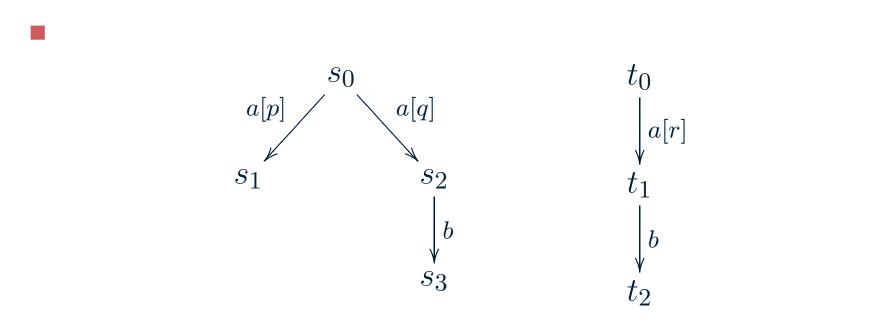
## But it is!



+

<

## But it is!



We add probabilities to the transitions.

- If p + q < r or p + q > r we can easily distinguish them.
- If p + q = r and p > 0 then q < r so  $\langle a \rangle_r \langle b \rangle_1 \top$  distinguishes them.

+

<

Show that the relation "s and s' satisfy exactly the same formulas" is a bisimulation.



<

- Show that the relation "s and s' satisfy exactly the same formulas" is a bisimulation.
- Can easily show that  $\tau_a(s, A) = \tau_a(s', A)$  for A of the form  $\llbracket \phi \rrbracket$ .



- Show that the relation "s and s' satisfy exactly the same formulas" is a bisimulation.
- Can easily show that  $\tau_a(s, A) = \tau_a(s', A)$  for A of the form  $\llbracket \phi \rrbracket$ .
- Use Dynkin's lemma to show that we get a well defined measure on the σ-algebra generated by such sets and the above equality holds.



+

<

- Show that the relation "s and s' satisfy exactly the same formulas" is a bisimulation.
- Can easily show that  $\tau_a(s, A) = \tau_a(s', A)$  for A of the form  $\llbracket \phi \rrbracket$ .
- Use Dynkin's lemma to show that we get a well defined measure on the σ-algebra generated by such sets and the above equality holds.
- Use special properties of analytic spaces to show that this σ-algebra is the same as the original σ-algebra.

>

## Simulation\*

Let  $S = (S, \Sigma, \tau)$  be a labelled Markov process. A preorder R on S is a **simulation** if whenever sRs', we have that for all  $a \in A$  and every R-closed measurable set  $A \in \Sigma$ ,  $\tau_a(s, A) \leq \tau_a(s', A)$ . We say s is simulated by s'if sRs' for some simulation relation R.



The logic used in the characterization has no negation, not even a limited negative construct.



- The logic used in the characterization has no negation, not even a limited negative construct.
- One can show that if s simulates s' then s satisfies all the formulas of L that s' satisfies.

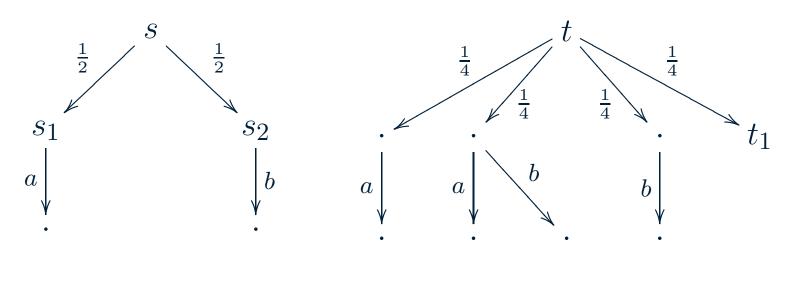




- The logic used in the characterization has no negation, not even a limited negative construct.
- One can show that if s simulates s' then s satisfies all the formulas of L that s' satisfies.
- What about the converse?



In the following picture, t satisfies all formulas of  $\mathcal{L}$  that s satisfies but t does not simulate s.



+

<

>

All transitions from s and t are labelled by a.

## **Counter Example (contd.)**

## • A formula of $\mathcal{L}$ that is satisfied by t but not by s.

# $\langle a \rangle_0 (\langle a \rangle_0 \mathsf{T} \land \langle b \rangle_0 \mathsf{T}).$



## **Counter Example (contd.)**

• A formula of  $\mathcal{L}$  that is satisfied by t but not by s.

 $\langle a \rangle_0 (\langle a \rangle_0 \mathsf{T} \land \langle b \rangle_0 \mathsf{T}).$ 

A formula with disjunction that is satisfied by s but not by t:

 $\langle a \rangle_{\frac{3}{4}} (\langle a \rangle_0 \mathsf{T} \lor \langle b \rangle_0 \mathsf{T}).$ 



+

## **A Logical Characterization for Simulation**

The logic *L* does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_{\vee} := \mathcal{L} \mid \phi_1 \lor \phi_2.$$



## **A Logical Characterization for Simulation**

The logic *L* does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_{\vee} := \mathcal{L} \mid \phi_1 \lor \phi_2.$$

With this logic we have: An LMP s<sub>1</sub> simulates s<sub>2</sub> if and only if for every formula \u03c6 of \u03c6<sub>\u2255</sub> we have

$$s_1 \models \phi \Rightarrow s_2 \models \phi.$$

<

>

+

## **A Logical Characterization for Simulation**

The logic *L* does **not** characterize simulation. One needs disjunction.

$$\mathcal{L}_{\vee} := \mathcal{L} \mid \phi_1 \lor \phi_2.$$

With this logic we have:
 An LMP s<sub>1</sub> simulates s<sub>2</sub> if and only if for every formula \$\phi\$ of \$\mathcal{L}\_V\$ we have

$$s_1 \models \phi \Rightarrow s_2 \models \phi.$$

<

>

The only proof we know uses domain theory.

+

#### Are exact equivalences reasonable??

Exact reasoning is unstable under (small) perturbations of probability numbers.



#### Are exact equivalences reasonable??

- Exact reasoning is unstable under (small) perturbations of probability numbers.
- Probability numbers are to be viewed as coming with some error estimate: reasoning principles based on the exact value of numbers are of dubious value.



#### Are exact equivalences reasonable??

- Exact reasoning is unstable under (small) perturbations of probability numbers.
- Probability numbers are to be viewed as coming with some error estimate: reasoning principles based on the exact value of numbers are of dubious value.
- Probability arises in the modelling of physical systems as an abstraction to specify incomplete knowledge.

<

### Are exact equivalences reasonable??

- Exact reasoning is unstable under (small) perturbations of probability numbers.
- Probability numbers are to be viewed as coming with some error estimate: reasoning principles based on the exact value of numbers are of dubious value.
- Probability arises in the modelling of physical systems as an abstraction to specify incomplete knowledge.
- Approximation of probability distributions is often used; e.g. in Monte Carlo schemes.

+

>

### A metric-based approximate viewpoint

Move from equality between processes to distances between processes (Jou and Smolka).



### A metric-based approximate viewpoint

- Move from equality between processes to distances between processes (Jou and Smolka).
- Formalize distance as a metric:

 $d(s,s) = 0, d(s,t) = d(t,s), d(s,u) \le d(s,t) + d(t,u)$ 



>

### A metric-based approximate viewpoint

- Move from equality between processes to distances between processes (Jou and Smolka).
- Formalize distance as a metric:

 $d(s,s) = 0, d(s,t) = d(t,s), d(s,u) \le d(s,t) + d(t,u)$ 

 Quantitative measurement of the distinction between processes.



#### Soundness:

 $d(s,t) = 0 \Leftrightarrow s, t \text{ are bisimilar}$ 



#### Soundness:

$$d(s,t) = 0 \Leftrightarrow s, t \text{ are bisimilar}$$

Stability of distance under temporal evolution: "Nearby states stay close forever."



<

Results work for Markov chains, Labelled Markov processes, Markov decision processes and Labelled Concurrent Markov chains with *τ*-transitions.



- Results work for Markov chains, Labelled Markov processes, Markov decision processes and Labelled Concurrent Markov chains with *τ*-transitions.
- Establishing closeness of states: Coinduction.



- Results work for Markov chains, Labelled Markov processes, Markov decision processes and Labelled Concurrent Markov chains with *τ*-transitions.
- Establishing closeness of states: Coinduction.
- Distinguishing states: Real-valued modal logics.

- Results work for Markov chains, Labelled Markov processes, Markov decision processes and Labelled Concurrent Markov chains with *τ*-transitions.
- Establishing closeness of states: Coinduction.
- Distinguishing states: Real-valued modal logics.

+

<

>

Equational and logical views coincide: Metrics yield same distances as real-valued modal logics.

- Results work for Markov chains, Labelled Markov processes, Markov decision processes and Labelled Concurrent Markov chains with *τ*-transitions.
- Establishing closeness of states: Coinduction.
- Distinguishing states: Real-valued modal logics.
- Equational and logical views coincide: Metrics yield same distances as real-valued modal logics.
- Compositional reasoning by Non-Expansivity. Process-combinators take closeby processes to closeby processes.

+

<

### **Approximation Results**

Our main result is a systematic approximation scheme for labeled Markov processes. The set of LMPs is a Polish space. Furthermore, our approximation results allow us to approximate integrals of continuous functions by computing them on finite approximants.





# **Approximation Results**

- Our main result is a systematic approximation scheme for labeled Markov processes. The set of LMPs is a Polish space. Furthermore, our approximation results allow us to approximate integrals of continuous functions by computing them on finite approximants.
- For any LMP, we explicitly provide a (countable) sequence of approximants to it such that:
  - For every logical property satisfied by a process, there is an element of the chain that also satisfies the property.

<

>

- The sequence of approximants converges – in a certain metric – to the process that is being approximated.

we establish the following equivalence of categories:

 $LMP \simeq Proc$ 

where **LMP** is the category with objects LMPs and with morphisms simulations; and *Proc* is the solution to the recursive domain equation

$$Proc \simeq \prod_{\text{Labels}} \mathcal{P}_{JP}(Proc).$$

<

>

we establish the following equivalence of categories:

 $LMP \simeq Proc$ 

where **LMP** is the category with objects LMPs and with morphisms simulations; and *Proc* is the solution to the recursive domain equation

$$Proc \simeq \prod_{\text{Labels}} \mathcal{P}_{JP}(Proc).$$

>

We show that there is a perfect match between:

- bisimulation and equality in Proc,
- simulation and the partial order of Proc,
- strict simulation and way below in *Proc*.

#### **Consequences of the Domain results\***

The sequence of approximants is a directed set in the simulation ordering and the process being approximated is the sup of this directed set.



### **Consequences of the Domain results\***

- The sequence of approximants is a directed set in the simulation ordering and the process being approximated is the sup of this directed set.
- The equivalence endows LMP with least upper bounds of ω-chains (wrt the simulation ordering). This shows that LMP can be used as the target of interpretation of a syntax that includes recursion.

>

+

We would like to "orient" the approximation process so that they are tailored to some formulas of interest.



- We would like to "orient" the approximation process so that they are tailored to some formulas of interest.
- We do not get the original process in the limit, but a bisimulation equivalent of it. Sometimes this is "spectacularly" not what we want.



- We would like to "orient" the approximation process so that they are tailored to some formulas of interest.
- We do not get the original process in the limit, but a bisimulation equivalent of it. Sometimes this is "spectacularly" not what we want.
- We can fix both the problems above but then we end up with the situation that the approximants are not LMPs. [DD, LICS03]

<

- We would like to "orient" the approximation process so that they are tailored to some formulas of interest.
- We do not get the original process in the limit, but a bisimulation equivalent of it. Sometimes this is "spectacularly" not what we want.
- We can fix both the problems above but then we end up with the situation that the approximants are not LMPs. [DD, LICS03]
- We can fix this too with a new approach to approximation based on conditional expectations. [DDP, CONCUR03]

+

<

We have a rich theory of LMPs with lots of possible directions for mathematical investigations.



- We have a rich theory of LMPs with lots of possible directions for mathematical investigations.
- We have also done weak bisimulation (probability mixed with nondeterminism), a metric for weak bisimulation and proved a logical characterization for weak bisimulation in terms of pCTL\*.



<

- We have a rich theory of LMPs with lots of possible directions for mathematical investigations.
- We have also done weak bisimulation (probability mixed with nondeterminism), a metric for weak bisimulation and proved a logical characterization for weak bisimulation in terms of pCTL\*.
- In early work on continuous time systems; we gave a logical characterization result for CTMCs [JALP 2003]. Later we showed how to define metrics for a very general class of systems (Generalized Semi-Markov Processes [QEST 2004].

+

<

- We have a rich theory of LMPs with lots of possible directions for mathematical investigations.
- We have also done weak bisimulation (probability mixed with nondeterminism), a metric for weak bisimulation and proved a logical characterization for weak bisimulation in terms of pCTL\*.
- In early work on continuous time systems; we gave a logical characterization result for CTMCs [JALP 2003]. Later we showed how to define metrics for a very general class of systems (Generalized Semi-Markov Processes [QEST 2004].
- In more recent work we have shown how to implement the approximation scheme using some Monte Carlo techniques [QEST 2005].