

Transition Systems and Bisimulations in Matrix Theory

Nikola Trčka

`n.trcka@tue.nl`

Eindhoven University of Technology

Faculty of Mathematics and Computer Science

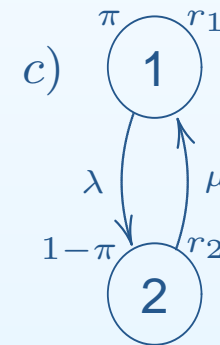
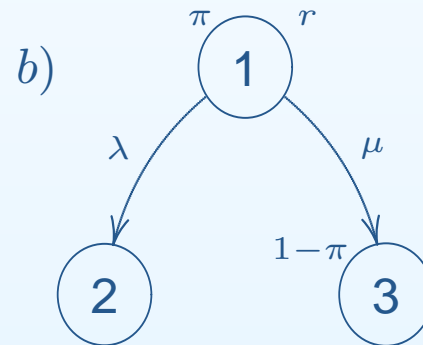
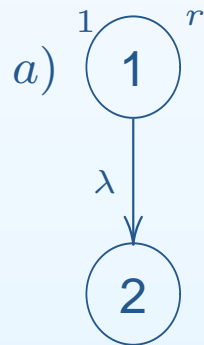
Formal Methods Group

Overview

- Markov reward chains
- Labeled transition systems
- Boolean matrix theory
- Transition systems as systems of boolean matrices
- Strong bisimulation = Ordinary lumping (Kemeny-Snell 1960)
- Weak bisimulation = τ -lumping (Markovski, Trčka, 2005)
- ? = τ -reduction (Doebelin 1938, Delebecque-Quadrat 1981, Coderch 1983)
- Branching bisimulation = ?

Continuous Time Markov Reward Chains

- Stochastic process with the history-less property (future independent of the past).
- Widely popular for quantitative analysis.
- Examples:

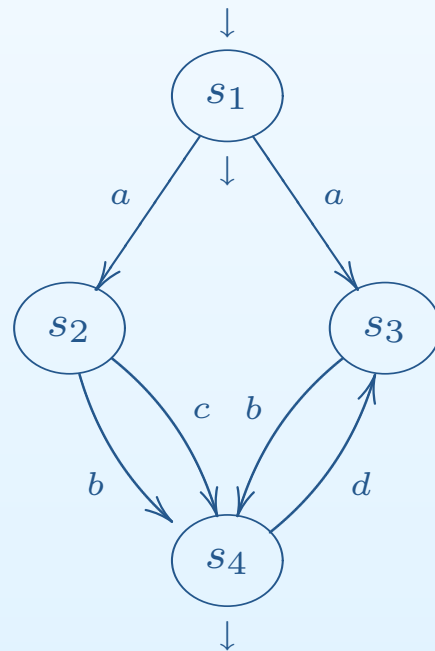


- Almost always explained in terms of matrix theory
- The Markov reward chain in c) is the triple (σ, Q, ρ) where

$$\sigma = \begin{pmatrix} \pi & 1-\pi \end{pmatrix}, \quad Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}, \quad \text{and} \quad \rho = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

Labeled Transition Systems

- Finite state
- One starting state
- Successful termination
- Example:



$A = \{a, b, c, d\}$ – set of actions

$\mathcal{S} = \{s_1, s_2, s_3, s_4\}$ – set of states

s_1 – initial state

s_1, s_4 – terminating states

Boolean Set algebra

- X – arbitrary set
- $\mathbb{P}(X) = (\mathcal{P}(X), \cup, \cap, ^-, \emptyset, X)$ – boolean algebra
- Different symbols (to emphasize connection with standard matrix theory):

$$\begin{aligned} + &\stackrel{\text{def}}{=} \cup \\ \cdot &\stackrel{\text{def}}{=} \cap \\ 0 &\stackrel{\text{def}}{=} \emptyset \\ 1 &\stackrel{\text{def}}{=} S. \end{aligned}$$

- Some axioms (for $\alpha \subseteq X$):

$$\alpha + \alpha = \alpha \cdot \alpha = \alpha$$

$$\alpha + 1 = 1$$

$$\alpha \cdot 1 = \alpha$$

$$\alpha + 0 = \alpha$$

$$\alpha \cdot 0 = 0.$$

- Order: $\alpha \leq \beta$ iff $\alpha + \beta = \beta$.

Boolean Matrix Theory

- Elements of $\mathbb{P}(X)^{n \times m}$
- Operations:

sum:	$(A + B)_{ij}$	$\stackrel{\text{def}}{=}$	$A_{ij} + B_{ij}$
product:	$(A \cdot B)_{ij}$	$\stackrel{\text{def}}{=}$	$\sum_k A_{ik} \cdot B_{kj}$
scalar product:	$(\alpha \cdot A)_{ij}$	$\stackrel{\text{def}}{=}$	$\alpha \cdot A_{ij}$
complement:	\bar{A}_{ij}	$\stackrel{\text{def}}{=}$	$\overline{A_{ij}}$
intersection:	$(A \sqcap B)_{ij}$	$\stackrel{\text{def}}{=}$	$A_{ij} \cdot B_{ij}$
transpose:	$(A^\top)_{ij}$	$\stackrel{\text{def}}{=}$	A_{ji}

- Order:

$$A \leq B \quad \text{iff} \quad A + B = B \quad \text{iff} \quad A_{ij} \leq B_{ij}$$

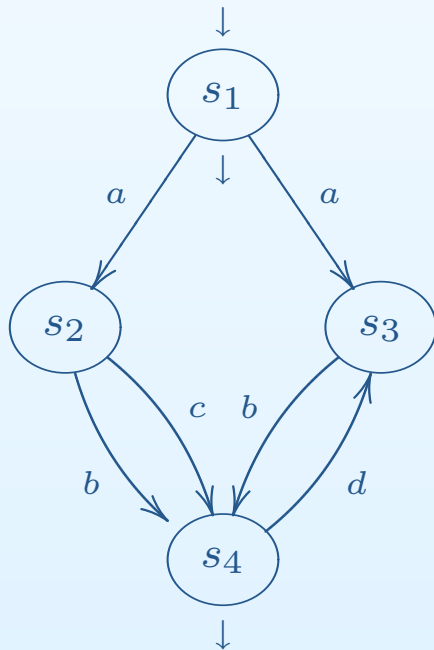
- Constants:

I – identity matrix, **0** – zero matrix

- 0–1 matrices = elements of $\{0, 1\}^{n \times m}$ (sometimes called relations)

Transition Systems as Systems of Matrices

- A – set of actions
- **DEF:** A transition system is a triple $\langle \sigma, A, \rho \rangle$ where:
 - $\sigma \in \{0, 1\}^{1 \times n}$ is the *initial vector*; its exactly one entry is 1,
 - $A \in \mathbb{P}(A)^{n \times n}$ is the *transition matrix*, and
 - $\rho \in \{0, 1\}^{n \times 1}$ is the *termination vector*.
- Example:



$$\sigma = (1 \ 0 \ 0 \ 0)$$

$$A = \begin{pmatrix} 0 & \{a\} & \{a\} & 0 \\ 0 & 0 & 0 & \{b, c\} \\ 0 & 0 & 0 & \{b\} \\ 0 & 0 & \{d\} & 0 \end{pmatrix} \quad \rho = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

Strong Bisimulation in Matrix Terms

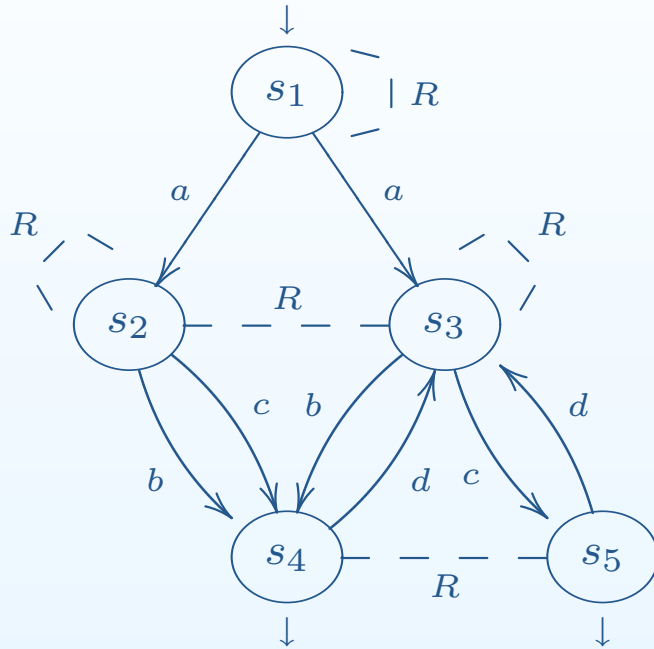
- Standard definition:
 - R – symmetric relation
 - $s_0 R s_0$ – initial state related to itself
 - **Termination condition:** if $s R t$ and $s \downarrow$, then $t \downarrow$
 - **Transfer condition:** if $s R t$ and $s \xrightarrow{a} s'$, then $t \xrightarrow{a} t'$ and $s' R t'$.
- Matrix definition:
 - $R \in \{0, 1\}^{n \times n}$, $R = R^T$
 - $\sigma \leq \sigma R$ – initial state related to itself
 - **Transfer condition:** $RA \leq AR$

$$\begin{array}{ccc}
 s_i - \frac{R}{\quad} - s_k & & s_i \\
 & \downarrow a & \downarrow a \\
 & s_j & s_\ell - \frac{R}{\quad} - s_j
 \end{array}
 \text{ implies}$$

- **Termination condition:** $R\rho \leq \rho$

$$s_i - \frac{R}{\quad} - s_j \downarrow \text{ implies } s_i \downarrow.$$

Strong Bisimulation – Example



$$\sigma = (1\ 0\ 0\ 0\ 0)$$

$$\rho = (0\ 0\ 0\ 1\ 1)^T$$

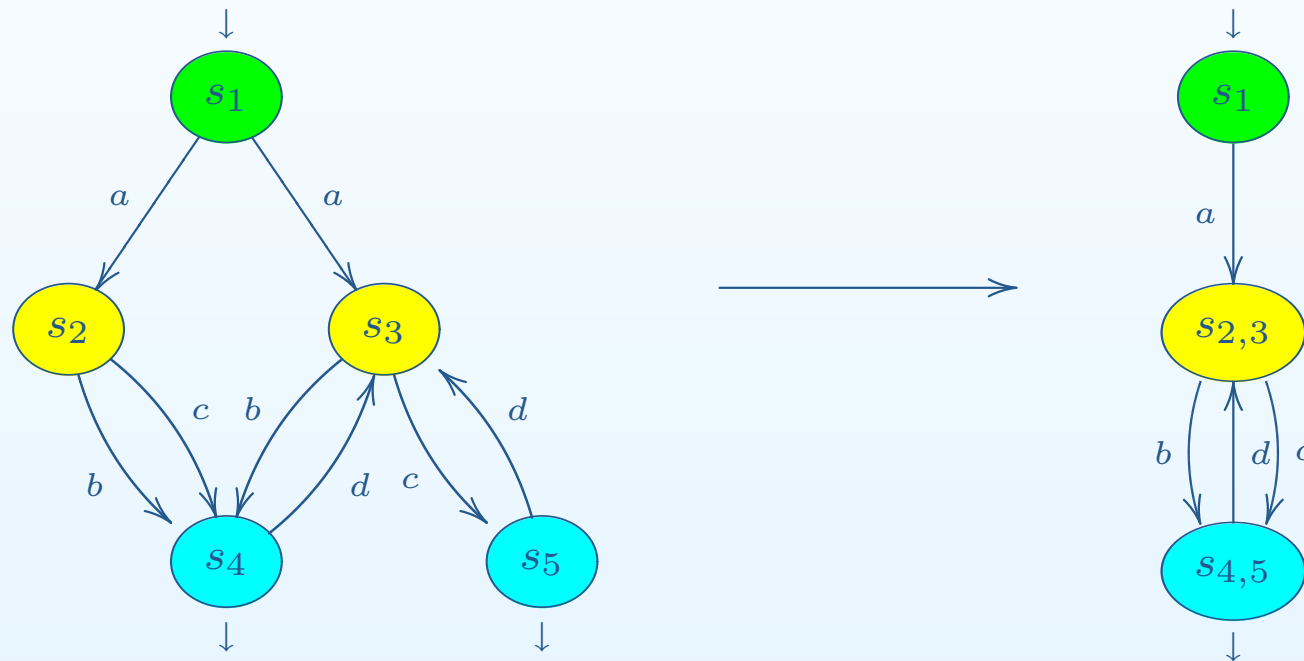
$$A = \begin{pmatrix} 0 & \{a\} & \{a\} & 0 & 0 \\ 0 & 0 & 0 & \{b,c\} & 0 \\ 0 & 0 & 0 & \{b\} & \{c\} \\ 0 & 0 & \{d\} & 0 & 0 \\ 0 & 0 & \{d\} & 0 & 0 \end{pmatrix}$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$RA = \begin{pmatrix} 0 & \{a\} & \{a\} & 0 & 0 \\ 0 & 0 & 0 & \{b,c\} & \{c\} \\ 0 & 0 & 0 & \{b,c\} & \{c\} \\ 0 & 0 & \{d\} & 0 & 0 \\ 0 & 0 & \{d\} & 0 & 0 \end{pmatrix} \leq \begin{pmatrix} 0 & \{a\} & \{a\} & 0 & 0 \\ 0 & 0 & 0 & \{b,c\} & \{b,c\} \\ 0 & 0 & 0 & \{b,c\} & \{b,c\} \\ 0 & \{d\} & \{d\} & 0 & 0 \\ 0 & \{d\} & \{d\} & 0 & 0 \end{pmatrix} = AR$$

Strong Lumping

- Aggregation modulo strong bisimulation equivalence
- Example:



Strong Lumping in Matrix Terms

- R is an equivalence relation: $I \leq R$, $R = R^T$, $R^2 \leq R$
- Unique decomposition: $R = VV^T$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} = VV^T$$

- $V \in \mathbb{R}^{n \times N}$ – called *collector* matrix
- $V^T V = I$
- Lumping condition is the requirement that $R = VV^T$ is a strong bisimulation:

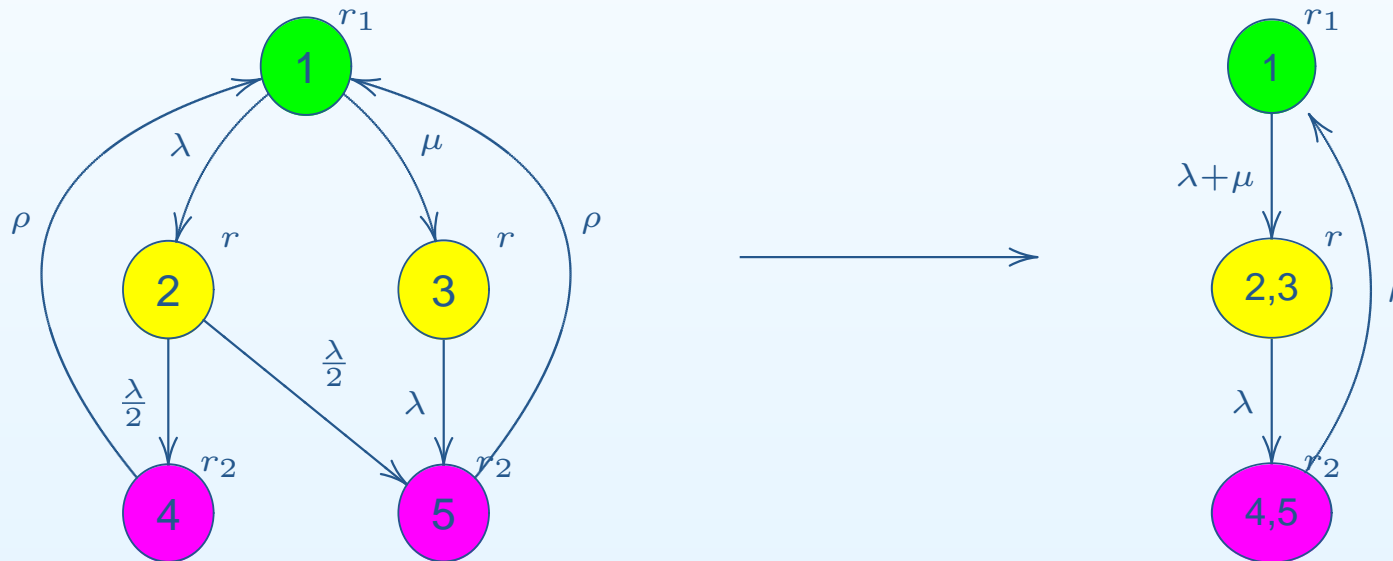
$$VUAV = AV \quad \text{and} \quad VU\rho = \rho, \quad \text{for any } U \text{ such that } UV = I$$

- No condition on σ because $\sigma = \sigma I \leq \sigma R$
- Lumped system $(\hat{\sigma}, \hat{A}, \hat{\rho})$ is defined by:

$$\hat{\sigma} = \sigma V, \quad \hat{A} = UAV \quad \text{and} \quad \hat{\rho} = U\rho$$

Strong Lumping = Ordinary Lumping for Markov chains

- Ordinary lumping was defined by Kemeny, Snell (1960)
- The most common way of aggregation
- All relevant properties preserved
- Example:



- Lumping condition:

$$VUQV = QV \quad \text{and} \quad VU\rho = \rho$$

- Lumped process:

$$\hat{Q} = UQV \quad \text{and} \quad \hat{\rho} = U\rho$$

Weak Bisimulation

- Standard definition:
 - **Termination condition:** if sRt and $s \downarrow$, then $t \Rightarrow t'$ and $t' \downarrow$
 - **Transfer condition:** if sRt and $s \xrightarrow{a} s'$, then $t \xRightarrow{a} t'$ and $s'Rt'$ (\xRightarrow{a} is $\Rightarrow \xrightarrow{a} \Rightarrow$).
- Matrix definition:
 - Unique representation of the transition matrix T
 - $T = A + \{\tau\} \cdot S$ where $S \in \{0, 1\}^{n \times n}$ and $\{\tau\} \cdot A = \mathbf{0}$
 - Example:

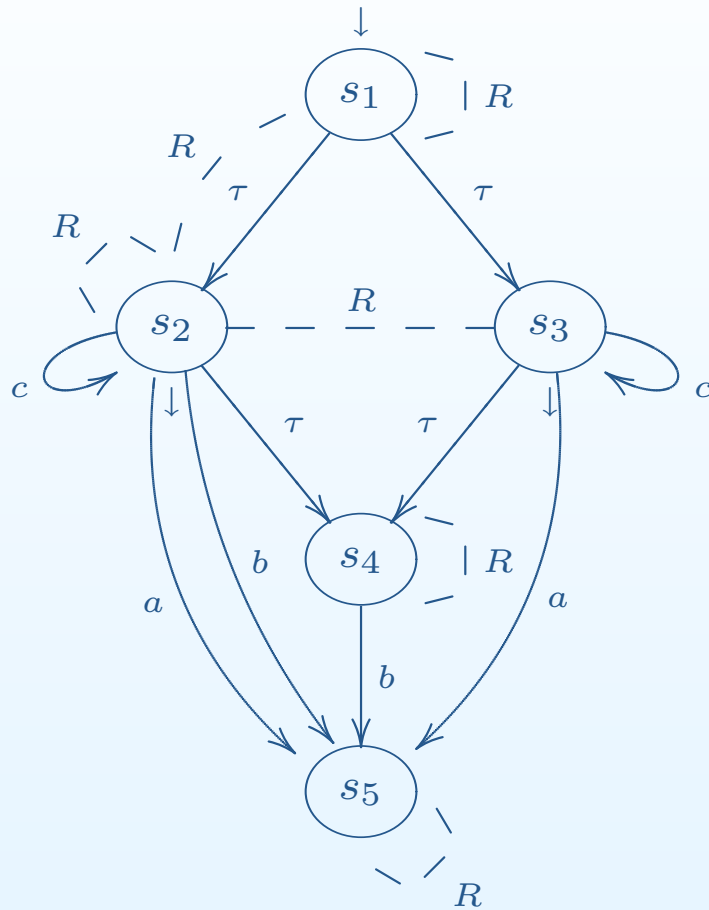
$$T = \begin{pmatrix} \{a\} & \{b, \tau\} \\ \{\tau\} & 0 \end{pmatrix} = \begin{pmatrix} \{a\} & \{b\} \\ 0 & 0 \end{pmatrix} + \{\tau\} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A + \{\tau\} \cdot S$$

- **Transfer condition for τ transitions:** $RS \leq S^*R$
- **Transfer condition for other actions:** $RA \leq S^*AS^*R$

$$\begin{array}{ccc}
 s_i & \xrightarrow{R} & s_k \\
 & & \downarrow a \\
 & & s_j
 \end{array}
 \quad \text{implies} \quad
 \begin{array}{ccc}
 s_i & & \\
 \Downarrow a & & \\
 s_\ell & \xrightarrow{R} & s_j
 \end{array}$$

- **Termination condition:** $R\rho \leq S^*\rho$

Weak Bisimulation – Example



$$\sigma = (1\ 0\ 0\ 0\ 0)$$

$$\rho = (0\ 1\ 1\ 0\ 0)^T$$

$$S = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \{a,b\} \\ 0 & 0 & \{c\} & 0 & \{a\} \\ 0 & 0 & 0 & \{c\} & \{b\} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

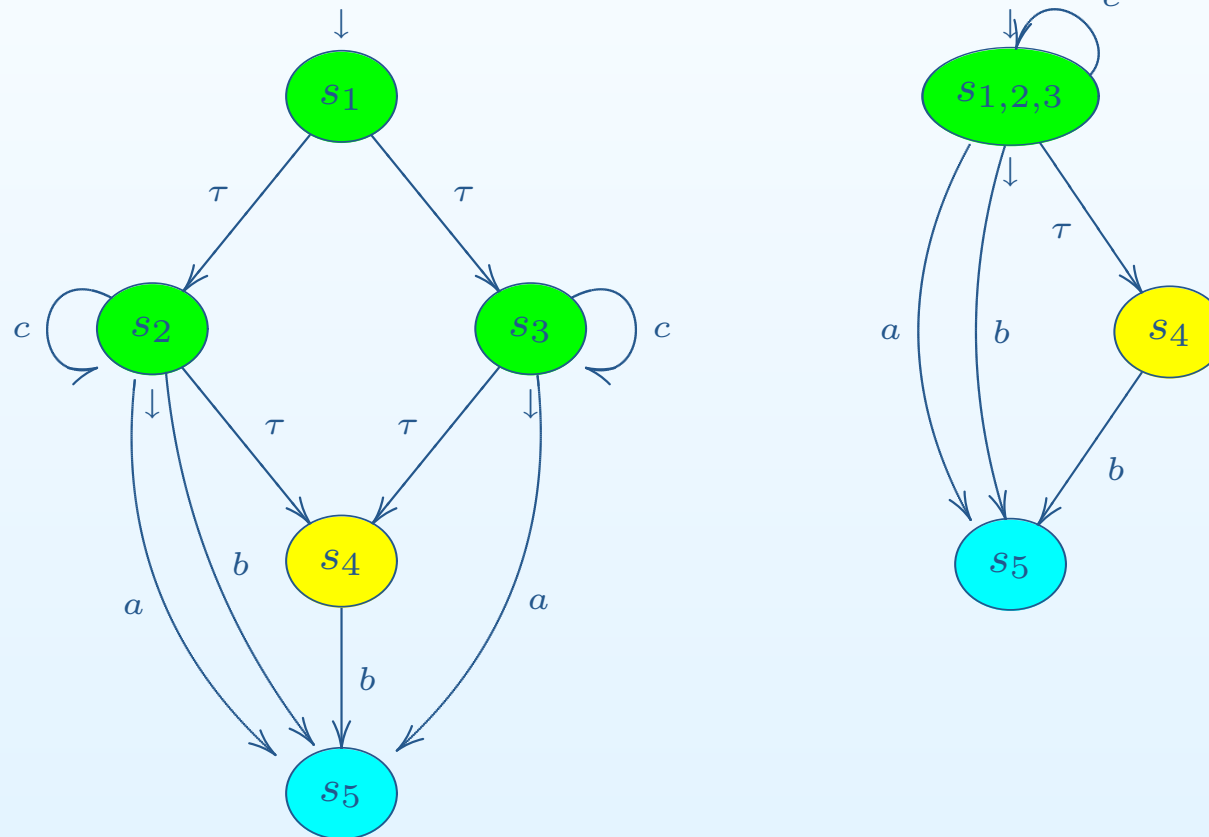
$$S^* = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad S^* A S^* = \begin{pmatrix} 0 & 0 & 0 & 0 & \{a,b\} \\ 0 & 0 & 0 & 0 & \{a,b\} \\ 0 & 0 & 0 & 0 & \{a,b\} \\ 0 & 0 & 0 & \{b\} & \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$R = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$RA = \begin{pmatrix} 0 & \{c\} & 0 & 0 & \{a,b\} \\ 0 & \{c\} & \{c\} & 0 & \{a,b\} \\ 0 & \{c\} & 0 & 0 & \{a,b\} \\ 0 & 0 & 0 & 0 & \{b\} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \leq \begin{pmatrix} \{c\} & \{c\} & \{c\} & \{c\} & \{a,b\} \\ \{c\} & \{c\} & \{c\} & \{c\} & \{a,b\} \\ 0 & \{c\} & 0 & \{c\} & \{a,b\} \\ 0 & 0 & 0 & 0 & \{b\} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = S^* A S^* R.$$

Weak Lumping

- Aggregation modulo weak bisimulation equivalence
- Example:



Weak Lumping in Matrix Terms

- $I \leq R, R = R^T, R^2 \leq R$
- Decomposition: $R = VV^T$
- Lumping condition is the requirement that $R = VV^T$ is a weak bisimulation:

$$VUS^*V = S^*V, \quad VUS^*AS^*V = S^*AS^*V, \quad \text{and} \quad VUS^*\rho = S^*\rho \quad (UV = I)$$

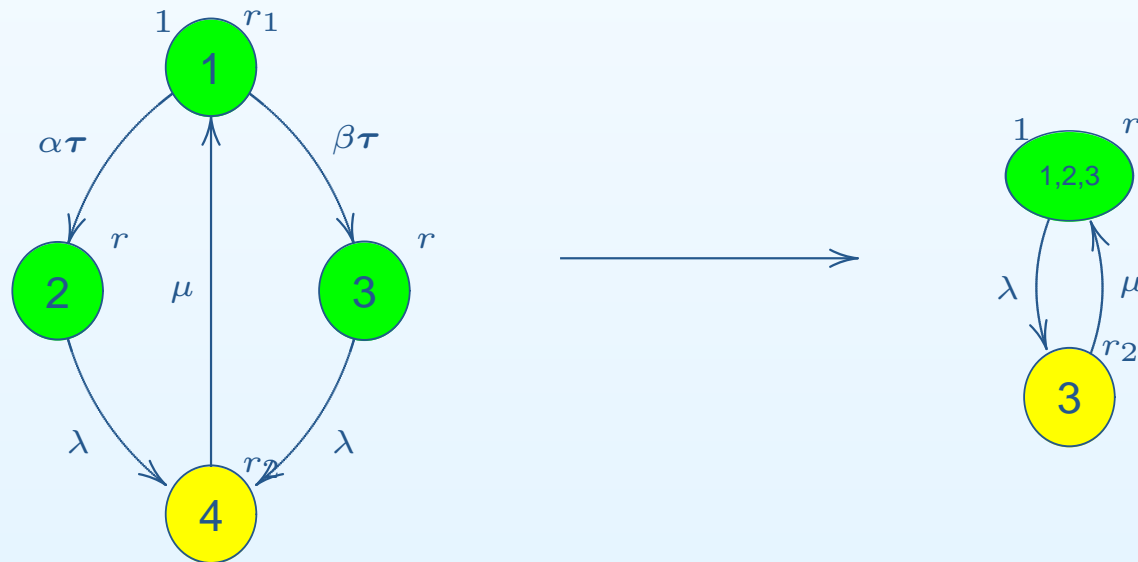
- No condition on σ because $\sigma = \sigma I \leq \sigma R$
- Lumped system is defined by:

$$\hat{\sigma} = \sigma V, \quad \hat{S} = V^T S V, \quad \hat{A} = V^T A V \quad \text{and} \quad \hat{\rho} = V^T \rho$$

- We cannot use arbitrary U but only V^T

τ -lumping for Markov Reward Chains

- Introduced by Markovski, Trčka (2005)
- Lumping method that abstracts from unobservable behavior
- Unobservable behavior = Instantaneous probabilistic behavior
- Example:



τ -lumping = Weak Bisimulation

- Follows from matrix representation of τ -lumping
- Unique decomposition of the generator: $Q = Q_s + \tau \cdot Q_f$
 - Q_s corresponds to A
 - Q_f corresponds to S
- $\Pi = \lim_{t \rightarrow \infty} e^{Q_f t}$ – ergodic projection
 - Π corresponds to S^* (they satisfy the same properties)

- Lumping condition:

$$VU\Pi V = \Pi V, \quad VU\Pi Q_s \Pi V = \Pi Q_s \Pi V, \quad \text{and} \quad VU\Pi \rho = \Pi \rho$$

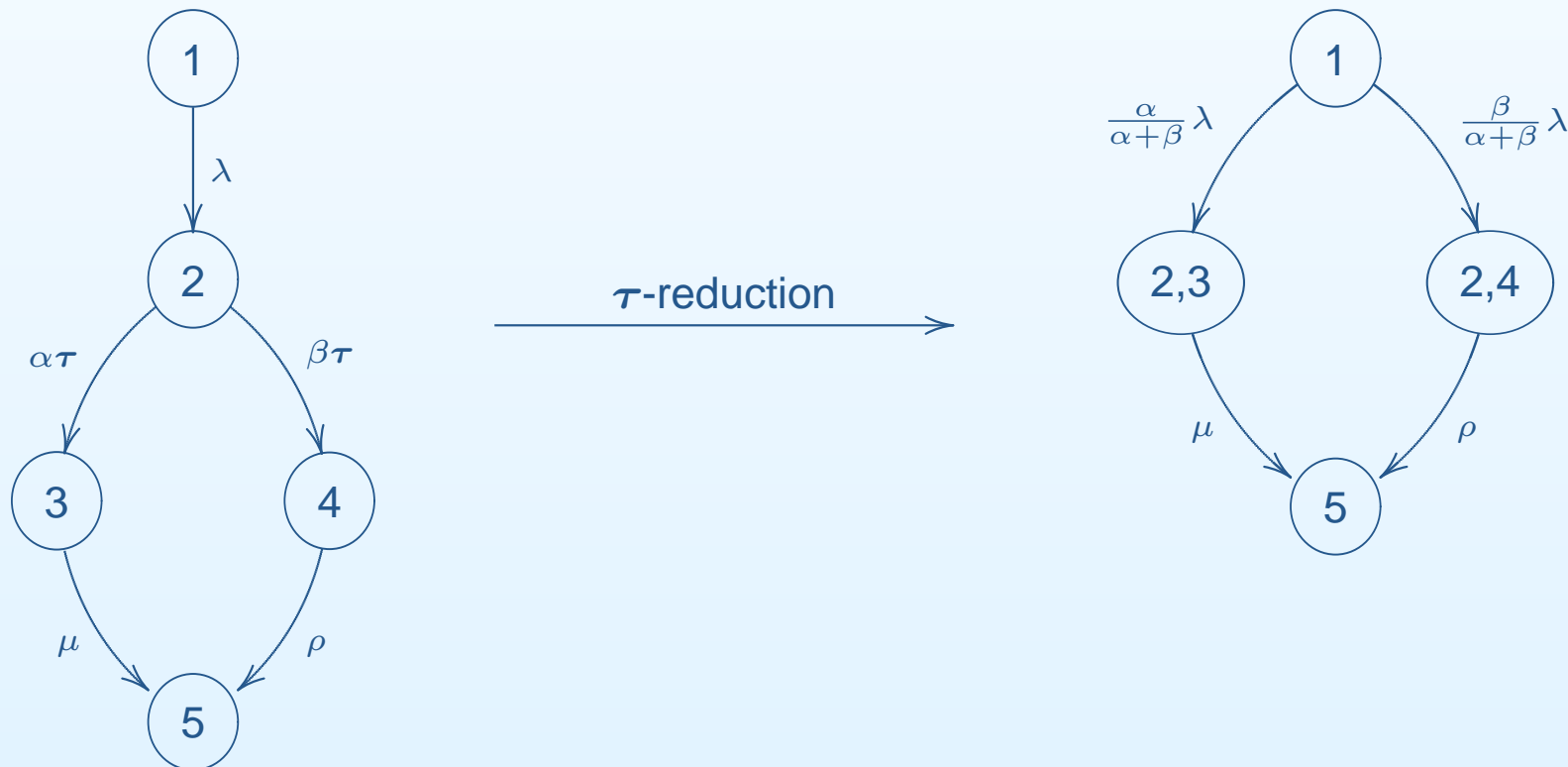
- Lumped process:

$$\hat{Q}_s = W Q_s V, \quad \hat{Q}_f = W Q_f V, \quad \text{and} \quad \hat{\rho} = W \rho \quad (W - \text{special})$$

- W corresponds to V^T (they satisfy the same properties)

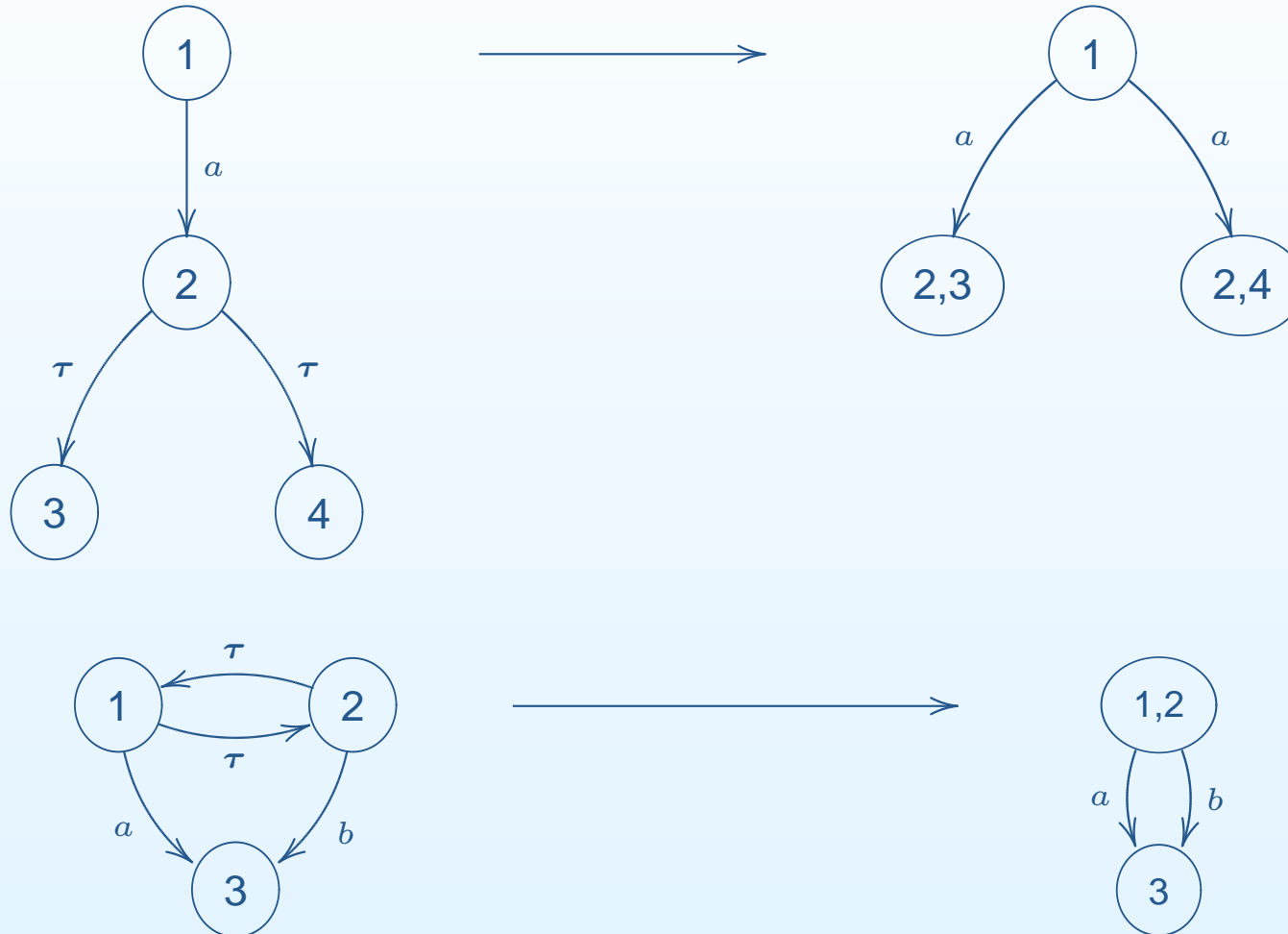
τ -reduction for Markov Chains

- Delebecque-Quadrat (1981), Coderch (1983); has roots in works of Doeblin (1938)
- Abstracts from unobservable behavior
- Can split states
- Example:



τ -reduction = ? for Transition Systems

- Matrix definition gives us:



Branching bisimulation

- Standard definition:
 - **Termination condition:** if sRt and $s \downarrow$, then $t \Rightarrow t'$, sRt' , and $t' \downarrow$
 - **Transfer condition:** if sRt and $s \xrightarrow{a} s'$, then $t \Rightarrow t'' \xrightarrow{a} t'$, sRt'' and $s'Rt'$
 - Cannot be directly expressed in matrix terms. Alternative: tRt'' instead of sRt''
- Matrix definition:
 - $T = A + \{\tau\} \cdot S$ where $S \in \{0, 1\}^{n \times n}$ and $\{\tau\} \cdot A = \mathbf{0}$
 - **Transfer condition for τ transitions:** $RS \leq (S^* \sqcap R)(I + S)R$
 - **Transfer condition for other actions:** $RA \leq (S^* \sqcap R)AR$



- **Termination condition:** $R\rho \leq (S^* \sqcap R)\rho$

Branching bisimulation = ? for Markov chains

- Problem: How to interpret \sqsubseteq
- Π is stochastic matrix but $\Pi \sqsubseteq R$ is not. Must be normalized somehow.
- Does it make sense?
- Is it stronger than τ -lumping?