Stochastic models in product form: the (E)RCAT methodology

Andrea Marin [1] Maria Grazia Vigliotti[2]

[1]Dipartimento di Informatica
Università Ca' Foscari di Venezia

[2]Department of Computing
Imperial College London

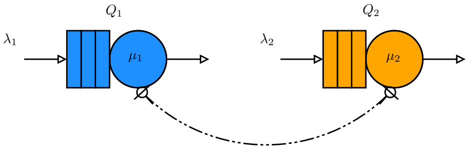
Second part: sketch

1. Introduction to the Extended Reversed Compound Agent Theorem (ERCAT) [Harrison '04a]
2. Applications for cooperations of pairs of automata which do not yield structural conditions of RCAT are shown
3. Special attention is devoted to queueing networks with finite capacity and blocking

Part II

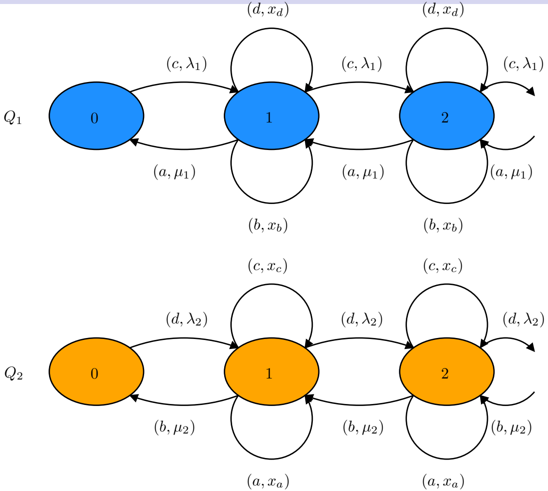
Extended Reversed Compound Agent Theorem (ERCAT)

A system in Boucherie's product-form



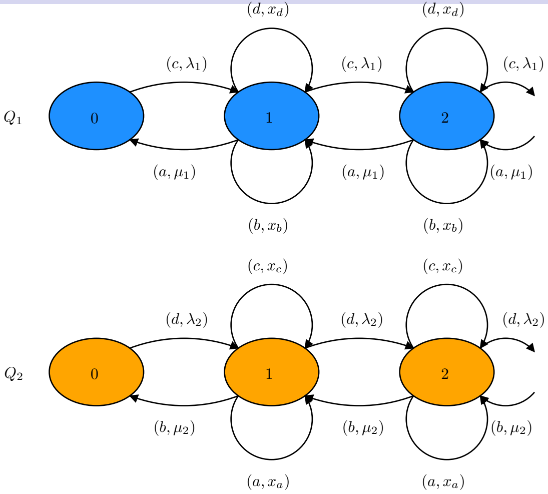
- Two exponential queues Q_1 and Q_2 with independent Poisson arrival streams with rate λ_1 and λ_2
- Service rates are μ_1 and μ_2
- If one of the queues enters in state 0 the other one is blocked (i.e. no arrivals or service completions occur)
- The model is known to be in Boucherie's product-form

Process representation



Are (G)RCAT structural conditions satisfied? NO!

Process representation



Are (G)RCAT structural conditions satisfied? NO!

A remark about reversed rates

- Consider two states n_i and n_j of cooperating automaton S
- Assume that there exists an active transition from n_i to n_j labelled by a with rate γ and that in n_j this is the *only* incoming transition with this label
- (G)RCAT requires to compute the value of:

$$K_a = \frac{\pi(n_i)}{\pi(n_j)}\gamma$$

- K_a may be interpreted as the rate of the transition from n_j to n_i in the reversed process of S
- We refer to K_a as the reversed rate of transitions labelled by a

Joint state space

- ERCAT requires to check a rate equation for each state of the irreducible subset of the joint process
- Often, states can be opportunely clustered and hence the computation becomes feasible
- The computational complexity is higher than the standard (G)RCAT
- Let (s_1, s_2) be a state of the irreducible subset of the joint process

Fundamental definitions

- $\mathcal{P}^{(s_1,s_2)\rightarrow}$: outgoing passive labels from s_1 or s_2
- $\mathcal{P}^{(s_1,s_2)\leftarrow}$: incoming passive labels into s_1 or s_2
- $\mathcal{A}^{(s_1,s_2)\rightarrow}$: outgoing active labels from s_1 or s_2
- $\mathcal{A}^{(s_1,s_2)\leftarrow}$: incoming active labels into s_1 or s_2
- $\alpha^{(s_1,s_2)}(a)$: rate of active transition labelled by a outgoing from (s_1, s_2)
- $\overline{\beta}^{(s_1,s_2)}(a)$: reversed rate of the passive transition labelled by a incoming into (s_1, s_2)

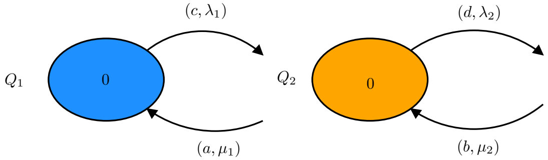
ERCAT formulation

Theorem (ERCAT)

Given two models Q_1 and Q_2 in which RCAT structural conditions are not satisfied but the reversed rates of the active transitions are constant, their cooperation is in product-form if the following rate equation is satisfied for each state (s_1, s_2) of the irreducible subset of states of the joint process:

$$\sum_{a \in \mathcal{P}^{(s_1,s_2)\rightarrow}} x_a - \sum_{a \in \mathcal{A}^{(s_1,s_2)\leftarrow}} x_a = \sum_{a \in \mathcal{P}^{(s_1,s_2)\leftarrow} \setminus \mathcal{A}^{(s_1,s_2)\leftarrow}} \overline{\beta}_a^{(s_1,s_2)} - \sum_{a \in \mathcal{A}^{(s_1,s_2)\rightarrow} \setminus \mathcal{P}^{(s_1,s_2)\rightarrow}} \alpha_a^{(s_1,s_2)}$$

State (0,0)



$$\mathcal{P}^{(0,0)\rightarrow} = \{\} \quad \mathcal{A}^{(0,0)\leftarrow} = \{a, b\}$$

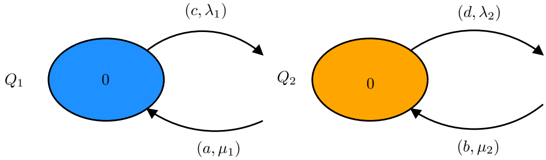
$$\mathcal{P}^{(0,0)\leftarrow} \setminus \mathcal{A}^{(0,0)\leftarrow} = \{\} \quad \mathcal{A}^{(0,0)\rightarrow} \setminus \mathcal{P}^{(0,0)\rightarrow} = \{c, d\}$$

$$-x_a - x_b = -\alpha_c^{(0,0)} - \alpha_d^{(0,0)}$$ Ok

Note that:

$$x_a = \lambda_1, \ \alpha_c^{(0,0)} = \lambda_1, \ x_b = \lambda_2, \alpha_d^{(0,0)} = \lambda_2$$

State (0,0)



$$\mathcal{P}^{(0,0)\rightarrow} = \{\} \quad \mathcal{A}^{(0,0)\leftarrow} = \{a, b\}$$

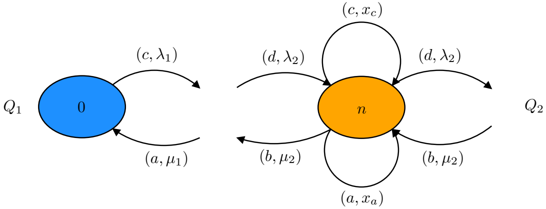
$$\mathcal{P}^{(0,0)\leftarrow} \setminus \mathcal{A}^{(0,0)\leftarrow} = \{\} \quad \mathcal{A}^{(0,0)\rightarrow} \setminus \mathcal{P}^{(0,0)\rightarrow} = \{c, d\}$$

$$-x_a - x_b = -\alpha_c^{(0,0)} - \alpha_d^{(0,0)} \quad \text{Ok}$$

Note that:

$$x_a = \lambda_1, \; \alpha_c^{(0,0)} = \lambda_1, \; x_b = \lambda_2, \alpha_d^{(0,0)} = \lambda_2$$

State (0,n), n>0



$$\mathcal{P}^{(0,n)\rightarrow} = \{a, c\} \quad \mathcal{A}^{(0,n)\leftarrow} = \{a, b, d\}$$

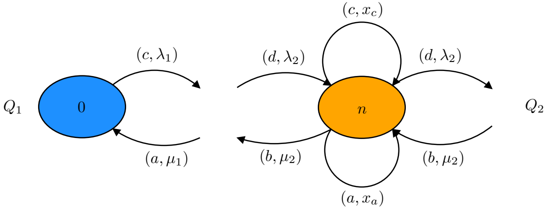
$$\mathcal{P}^{(0,n)\leftarrow} \setminus \mathcal{A}^{(0,n)\leftarrow} = \{c\} \quad \mathcal{A}^{(0,n)\rightarrow} \setminus \mathcal{P}^{(0,n)\rightarrow} = \{b, d\}$$

$$x_a + x_c - x_a - x_b - x_d = \overline{\beta}_c^{(0,n)} - \alpha_b^{(0,n)} - \alpha_d^{(0,n)}$$ Ok!

Note that:

$$x_b = \lambda_2,\ x_c = \mu_1,\ x_d = \mu_2,\ \overline{\beta}_c^{(0,n)} = \mu_1,\ \alpha_b^{(0,n)} = \mu_2,\ \alpha_d^{(0,n)} = \lambda_2$$

State (0,n), n>0



$$\mathcal{P}^{(0,n)\rightarrow} = \{a, c\} \quad \mathcal{A}^{(0,n)\leftarrow} = \{a, b, d\}$$

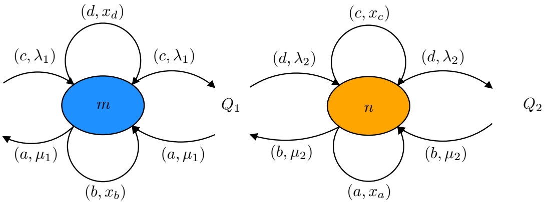
$$\mathcal{P}^{(0,n)\leftarrow} \setminus \mathcal{A}^{(0,n)\leftarrow} = \{c\} \quad \mathcal{A}^{(0,n)\rightarrow} \setminus \mathcal{P}^{(0,n)\rightarrow} = \{b, d\}$$

$$x_a + x_c - x_a - x_b - x_d = \overline{\beta}_c^{(0,n)} - \alpha_b^{(0,n)} - \alpha_d^{(0,n)}$$ Ok!

Note that:

$$x_b = \lambda_2,\ x_c = \mu_1,\ x_d = \mu_2,\ \overline{\beta}_c^{(0,n)} = \mu_1,\ \alpha_b^{(0,n)} = \mu_2,\ \alpha_d^{(0,n)} = \lambda_2$$

State (m,n), m,n>0



$$\mathcal{P}^{(m,n)\rightarrow} = \{a, b, c, d\} \quad \mathcal{A}^{(m,n)\leftarrow} = \{a, b, c, d\}$$
$$\mathcal{P}^{(m,n)\leftarrow} \setminus \mathcal{A}^{(m,n)\leftarrow} = \{\} \quad \mathcal{A}^{(m,n)\rightarrow} \setminus \mathcal{P}^{(m,n)\rightarrow} = \{\}$$

0=0

Note that states $(m, 0)$ with $m > 0$ are similar to $(0, n)$, $n > 0$.

Conclusion of the running example

- The model, as expected, is in product-form:

$$\pi(m,n) \propto \left(\frac{\lambda_1}{\mu_1}\right)^m \left(\frac{\lambda_2}{\mu_2}\right)^n$$

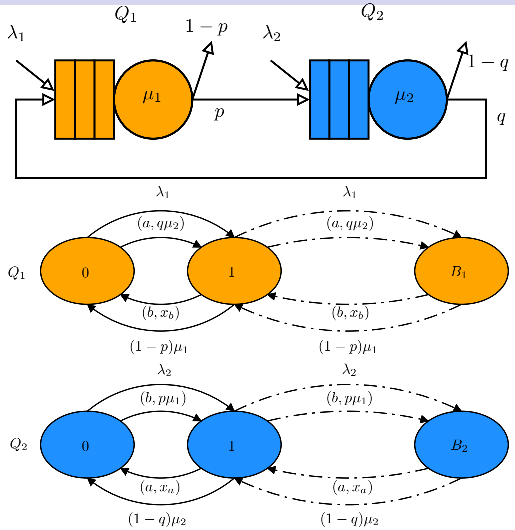
- Note that state $(0,0)$ is either the only ergodic state or does not belong to the irreducible subset
- Hence, the normalising constant distinguishes this solution from the case of independent queues
- Every Boucherie's product-form with full blocking can be studied by ERCAT [Harrison '04a]

Queues with finite capacity and Repetitive Service (RS) blocking

- We consider a network of queues, $Q_1, \ldots, Q_N$ with finite capacity B_i and service rate μ_i
- At a job completion at Q_i the customer goes to Q_j with probability P_{ij}. If Q_j if saturated the customer service is restarted and a new target station is selected at job completion
- In open networks λ_i is the arrival rate at Q_i and customers leave the system with probability $1 - \sum_j P_{ij}$. Arrivals at saturated queues are not allowed

Andrea Marin

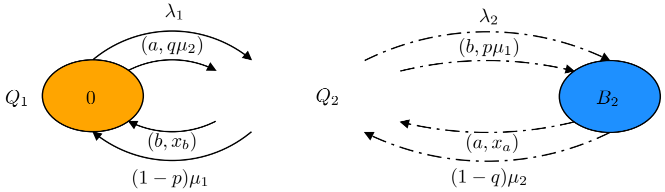
Example



Notes

- Differently from ordinary queueing networks we use active transitions to model synchronised arrivals and passive to model synchronised departures
- Which states shall we consider?
  1. $(0, 0)$
  2. $(0, K)$ with $0 < K < B_2$ (and symmetrically we obtain $(K, 0)$ with $0 < K < B_1$)
  3. $(0, B_2)$
  4. (K, B_2) with $0 < K < B_1$ (and symmetrically we obtain $(0, K)$ with $0 < K < B_2$)
  5. (B_1, B_2)
- Note that $\alpha_a^{(\cdot,\cdot)} = q\mu_2$, $\alpha_b^{(\cdot,\cdot)} = p\mu_1$ and also $\overline{\beta}_a^{(\cdot,\cdot)} = \overline{\beta}_a$ and $\overline{\beta}_b^{(\cdot,\cdot)} = \overline{\beta}_b$

State $(0, B_2)$



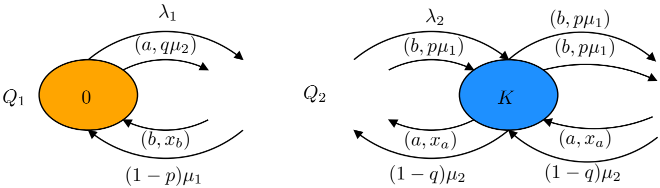
$$\mathcal{P}^{(0,B_2)\to} = \{a\} \quad \mathcal{A}^{(0,B_2)\leftarrow} = \{b\}$$

$$\mathcal{A}^{(0,B_2)\to} \setminus \mathcal{P}^{(0,B_2)\to} = \{\} \quad \mathcal{P}^{(0,B_2)\leftarrow} \setminus \mathcal{A}^{(0,B_2)\leftarrow} = \{\}$$

$$x_a - x_b = 0 \Rightarrow x_a = x_b \tag{1}$$

Andrea Marin

State $(0, K)$



$$\mathcal{P}^{(0,K)\rightarrow} = \{a\} \quad \mathcal{A}^{(0,K)\leftarrow} = \{b\}$$

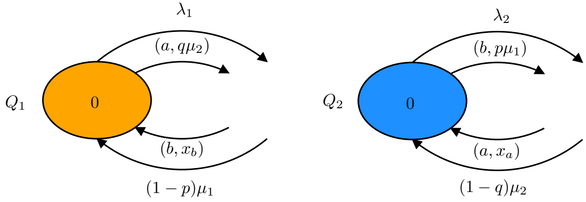
$$\mathcal{A}^{(0,K)\rightarrow} \setminus \mathcal{P}^{(0,K)\rightarrow} = \{b\} \quad \mathcal{P}^{(0,K)\leftarrow} \setminus \mathcal{A}^{(0,K)\leftarrow} = \{a\}$$

i.e.:

$$x_a - x_b = \overline{\beta}_b^{(0,K)} - \alpha_a^{(0,K)} \xrightarrow{(1)} \overline{\beta}_b = \alpha_a \qquad (2)$$

By symmetry, state $(K, 0)$ gives $\overline{\beta}_a = \alpha_b$

State $(0, 0)$



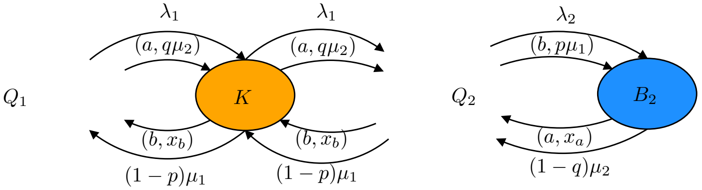
$$\mathcal{P}^{(0,0)\rightarrow} = \{\} \quad \mathcal{A}^{(0,0)\leftarrow} = \{\}$$

$$\mathcal{A}^{(0,0)\rightarrow} \setminus \mathcal{P}^{(0,0)\rightarrow} = \{a, b\} \quad \mathcal{P}^{(0,0)\leftarrow} \setminus \mathcal{A}^{(0,0)\leftarrow} = \{a, b\}$$

$$\overline{\beta}_a^{(0,0)} + \overline{\beta}_b^{(0,0)} = \alpha_a^{(0,0)} + \alpha_b^{(0,0)}$$

which is a consequence of (2)

States (K, B_2) and (B_1, K), (B_1, B_2)



- For these states we have:
  - $\mathcal{P}^{(\cdot,\cdot)\rightarrow} = \{a, b\}$
  - $\mathcal{A}^{(\cdot,\cdot)\leftarrow} = \{a, b\}$
- Since all the synchronising labels are present in both these sets, the rate equation for these states is an identity.

Conditions derived from the ERCAT rate equations

$$\begin{cases} x_a = x_b \\ \overline{\beta}_b = \alpha_a = q\mu_2 \\ \overline{\beta}_a = \alpha_b = p\mu_1 \end{cases}$$

The process analysis gives:

$$\overline{\beta}_b = \frac{x_b(\lambda_1 + q\mu_2)}{x_b + (1-p)\mu_1} \quad \overline{\beta}_a = \frac{x_a(\lambda_2 + p\mu_1)}{x_a + (1-q)\mu_2}$$

From which we straightforwardly derive:

$$x_a = \frac{(1-q)p\mu_1\mu_2}{\lambda_2} \quad x_b = \frac{(1-p)q\mu_1\mu_2}{\lambda_1} \tag{3}$$

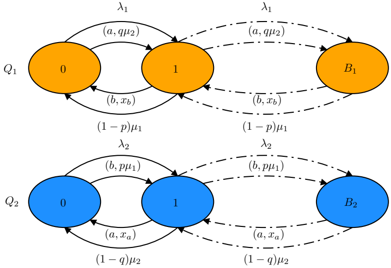
Product-form rate condition

Since $x_a = x_b$ by (1) we have the product-form rate condition:

$$(1-p)q\lambda_2 = (1-q)p\lambda_1$$

Under this assumption expressions (3) for x_a, x_b satisfies:

$$x_a = \frac{(x_b + (1-p)\mu_1)q\mu_2}{\lambda_1 + q\mu_2} \quad x_b = \frac{(x_a + (1-q)\mu_2)p\mu_1}{\lambda_2 + p\mu_1}$$



Generalisation

- ERCAT may be applied to a set of agent with pairwise cooperations (this is also known a MARCAT)
- In case of QN with RS blocking and general topology in [Balsamo et al. '10] is proved that:

Theorem

A QN (open or closed) with finite capacity stations and RS blocking policy with reversible routing matrix always satisfies ERCAT rate equations.

- Product-form for reversible routing has been proved in [Akyildiz '87]

Closed QN with RS blocking

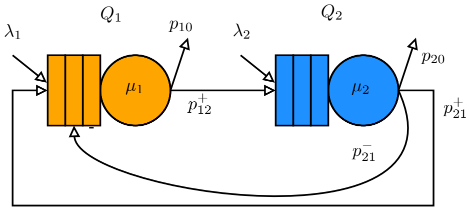
- Consider a closed QN with RS blocking policy
- Note that the ERCAT rate equation is an identity for state $\mathbf{n}$ when none of the stations is empty in $\mathbf{n}$
- We immediately have the following result:

Theorem (QN with strict non-empty condition)

A closed QN with finite capacity stations and RS blocking is in product-form if the number of customers is such that none of the station can be empty (strict non-empty condition)

- In [Balsamo et al. '10] we prove that the same result for QN in which at most one station can be empty (non-empty condition)

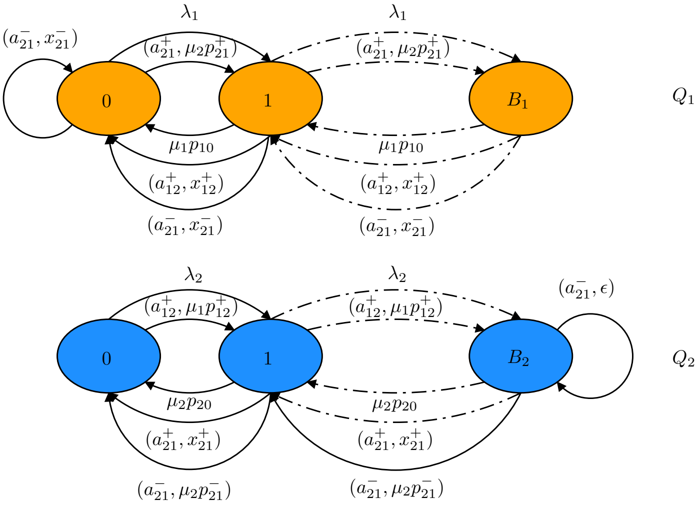
Model description



- $p_{10} + p_{12}^+ = 1$
- $p_{21}^+ + p_{21}^- + p_{20} = 1$

- Exp. service times with parameters μ_1 and μ_2
- Independent Poisson arrival processes
- A customer leaving Q_2 may:
  - leave the system with pr. p_{20}
  - enter Q_1 as a standard customer with pr. p_{21}^+
  - delete a customer in Q_1 with pr. p_{21}^-
- While Q_2 is saturated, customers from Q_1 are deleted with rate ϵ

Underlying processes



Key-idea

Observe that label a_{21}^{-} does not change ERCAT rate condition:

- a_{21}^{-} exits from every state of Q_1 (passive):

$$a_{21}^{-} \in \mathcal{P}^{(s_1,s_2)\rightarrow} \quad 0 \leq s_1 \leq B_1, 0 \leq s_2 \leq B_2$$

- a_{21}^{-} enters into every state of Q_2 (active):

$$a_{21}^{-} \in \mathcal{A}^{(s_1,s_2)\leftarrow} \quad 0 \leq s_1 \leq B_1, 0 \leq s_2 \leq B_2$$

ERCAT rate condition:

$$\sum_{a \in \mathcal{P}^{(s_1,s_2)\rightarrow}} x_a - \sum_{a \in \mathcal{A}^{(s_1,s_2)\leftarrow}} x_a$$
$$= \sum_{a \in \mathcal{P}^{(s_1,s_2)\leftarrow} \setminus \mathcal{A}^{(s_1,s_2)\leftarrow}} \overline{\beta}_a^{(s_1,s_2)} - \sum_{a \in \mathcal{A}^{(s_1,s_2)\rightarrow} \setminus \mathcal{P}^{(s_1,s_2)\rightarrow}} \alpha_a^{(s_1,s_2)}$$

Product-form conditions

We still have $x^+_{12} = x^+_{21}$, where:

$$
\begin{aligned}
x^+_{12} &= \frac{\mu_1 p^+_{12}}{\lambda_2 + \mu_1 p^+_{12}}(\mu_2 p_{20} + x^+_{21} + \mu_2 p^+_{21}) \\
x^+_{21} &= \frac{\mu_2 p^+_{21}}{\lambda_1 + \mu_2 p^+_{21}}(\mu_1 p_{10} + x^+_{12} + x^-_{21}) \\
x^-_{21} &= \frac{\mu_2 p^-_{21}}{\mu_2 p^-_{21} + x^+_{21} + \mu_2 p_{20}}(\lambda_2 + \mu_1 p^+_{12})
\end{aligned}
$$

After some algebra we derive the condition:

$$
\lambda_1 p^+_{12}(1 - p^-_{21}) = \lambda_2 p^+_{21}\left(p_{10} + \frac{\lambda_2}{\mu_1}\frac{p^-_{21}}{1 - p^+_{21}}\right)
$$

Explicit expressions of x_{12}^+, x_{21}^+, x_{21}^-, ϵ

- We derive the expressions for $x_{21}^+ = x_{12}^+$ and x_{21}^-:

$$x_{21}^+ = \frac{\mu_2 p_{21}^+}{\lambda_1}\left(\mu_1 p_{10} + \frac{\lambda_2 p_{21}^-}{1 - p_{21}^+}\right)$$

$$x_{21}^- = \frac{\lambda_2 p_{21}^-}{1 - p_{21}^+}$$

- Constant reverse rates of the active transitions:

$$\epsilon = x_{21}^-$$

Product-form expression

- Product-form expression:

$$\pi(n_1, n_2) \propto \left(\frac{\lambda_1 + \mu_2 p_{21}^+}{\mu_1 p_{10} + x_{12}^+ + x_{21}^-} \right)^{n_1} \cdot \left(\frac{\lambda_2 + \mu_1 p_{12}^+}{\mu_2 p_{20} + x_{21}^+ + \mu_2 p_{21}^-} \right)^{n_2}$$

for $0 \leq n_1 \leq B_1$ and $0 \leq n_2 \leq B_2$

Final remarks

- In the second part of the tutorial we have shown how to overcome some limitations of original RCAT and GRCAT formulation
  - In case structural conditions of (G)RCAT are not satisfy we may apply ERCAT
- Application of (G)RCAT or ERCAT may be done algorithmically, however the computational cost of ERCAT is higher than that of (G)RCAT

Applications

- Other models than those presented here may be studied by RCAT and its extensions (e.g. product-form Stochastic Petri Nets)
- New product-form may be derived
- The solution of the traffic equations may be efficiently computed by means of the algorithm presented in [Marin et al. '09]
  - Numerical and iterative algorithm
- Product-form of models expressed in terms of different formalisms may be derived.

Appendix: Reversible routing matrix

- Consider a queueing network with N stations and fixed routing probability matrix $\mathbf{P} = [p_{ij}]$, $1 \leq i, j \leq N$
- p_{i0} is the probability of leaving the network after a job completion at station i
- e_i is the (relative) visit ratio to station i
- λ_i is the arrival rate at station i

Definition (Reversible routing matrix)

The routing matrix **P** is said reversible if:

$$\begin{cases} e_i p_{ij} = e_j p_{ji} & \text{for } 1 \leq i, j \leq N \\ \lambda_i = e_i p_{i0} & \text{for } 1 \leq i \leq N \end{cases}$$

For Further Reading I

- B. Pittel:Closed exponential networks of queues with saturation: The Jackson-type stationary distribution and its asymptotic analysis,
  Math. of Op. Res., vol. 4, n. 4, pp. 357–378, 1979
- I.F. Akyildiz: Exact product form solution for queueing networks with blocking,
  IEEE Trans. on Computers, vol. C-36-1, pp. 122-125, 1987
- P.G. Harrison: Turning back time in Markovian process algebra,
  Theoretical Computer Science, vol. 290, n. 3, pp. 1947–1986, 2003

For Further Reading II

- P.G. Harrison: Reversed processes, product forms and a non-product form,
  Linear Algebra and its App., vol. 386, pp. 359–381, 2004.
- P.G. Harrison: Compositional reversed Markov processes, with applications to G-networks,
  Perf. Eval., vol. 57, n. 3, pp. 379–408, 2004
- A. Marin and S. R. Bulò: A general algorithm to compute the steady-state solution of product-form cooperating Markov chains,
  in Proc. of MASCOTS 2009, pp. 515–524, 2009

For Further Reading III

- A. Marin, M.G. Vigliotti: A general result for deriving product-form solutions of Markovian models,
  Proc. of WOSP/SIPEW Int. Conf. on Perf. Eval.
- S. Balsamo, P. G. Harrison, A. Marin: A unifying approach to product-forms in networks with finite capacity constraints,
  Proc. of the 2010 ACM SIGMETRICS, pp. 25–36, 2010