MARKOV REWARD MODELS (MRMS)

Presented by: Hadi Salimi
Dependable Software Design
Computer Engineering Department
IRAN University of Science and Technology
hsalimi@iust.ac.ir

INTRODUCTION

$$\frac{\mathrm{d}\,\boldsymbol{\pi}(t)}{\mathrm{d}\,t} = \boldsymbol{\pi}(t)\mathbf{Q}\,, \qquad \boldsymbol{\pi}(0) = \boldsymbol{\pi}_0$$

- For the sake of completeness, we repeat here the fundamental equations of CTMC analysis.

CUMULATIVE PROBABILITIES

- Sometimes cumulative probabilities are of interest, so we will have:

$$\mathbf{L}(t) = \int_0^t \boldsymbol{\pi}(u)\,\mathrm{d}\,u$$

TIME-AVERAGE BEHAVIOR

- Closely related to the vector of cumulative state probabilities is the vector describing the time-average behavior of the CTMC:

$$\mathbf{M}(t) = \frac{1}{t}\mathbf{L}(t)$$

MRMS

- MRMs have long been used in Markov decision theory to assign cost and reward structures to states of Markov processes for an optimization.
- Meyer adopted MRMs to provide a framework for an integrated approach to performance and dependability characteristics.
- He coined the term *performability to refer to measures characterizing the ability of fault-tolerant* systems, that is, systems that are subject to component failures and that can perform certain tasks in the presence of failures.

ASSIGNING REWARDS

- With MRMs, rewards can be assigned to states or to transitions between states of a CTMC.
- In the former case, these rewards are referred to as *reward rates and in the latter as impulse rewards*.
- *In this text we consider state-based* rewards only.

REWARDS

- The reward rates are defined based on the system requirements, be it availability-, reliability-, or task-oriented performance.
- Let the reward rate r_i, *be assigned* to state *i€S*.
- Then, a reward $r_i t_i$ *is accrued during a sojourn of time Ti in state i.*

REWARD RATES

- Let { X (t) , t >= 0} denote a homogeneous finite-state CTMC with state space S.
- Then, the random variable Z(t) refers to the instantaneous reward rate of the MRM at time t.

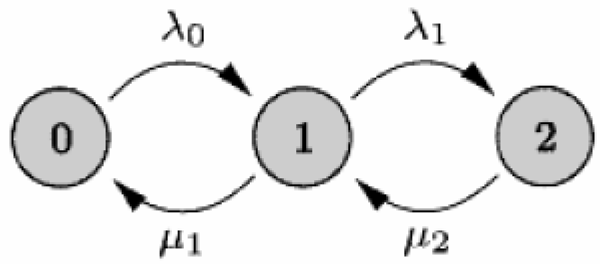
$$Z(t) = r_{X(t)}$$

ACCUMULATED REWARD

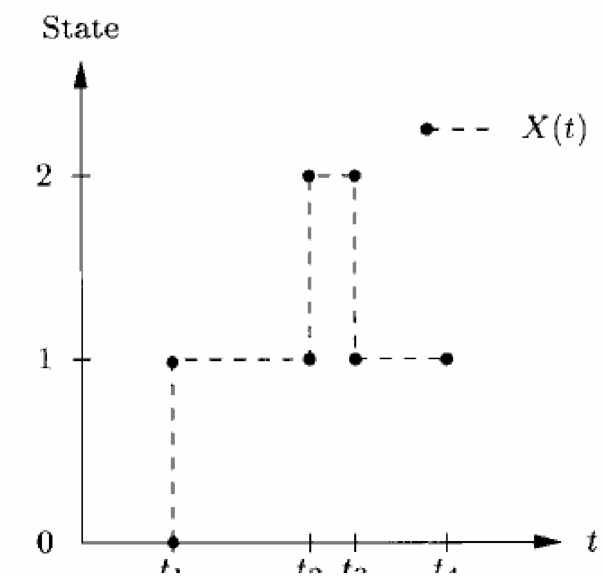
- The *accumulated reward* *Y (t) in the finite time horizon [0, t) is given by:*

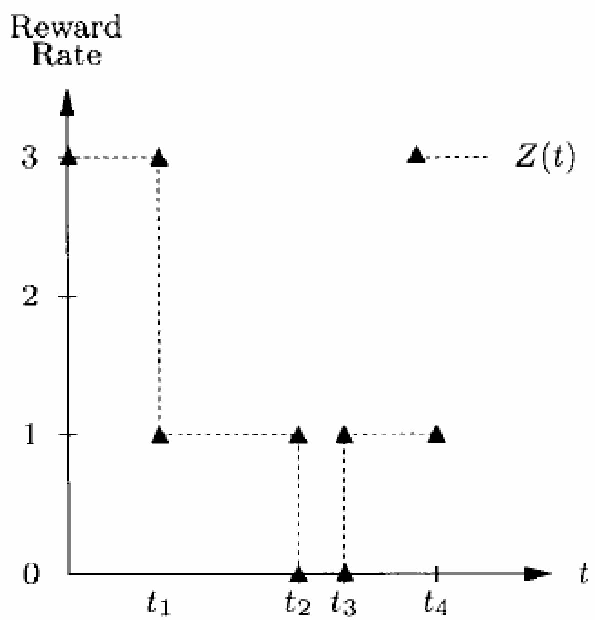
$$Y(t) = \int_0^t Z(\tau)\, \mathrm{d}\,\tau$$

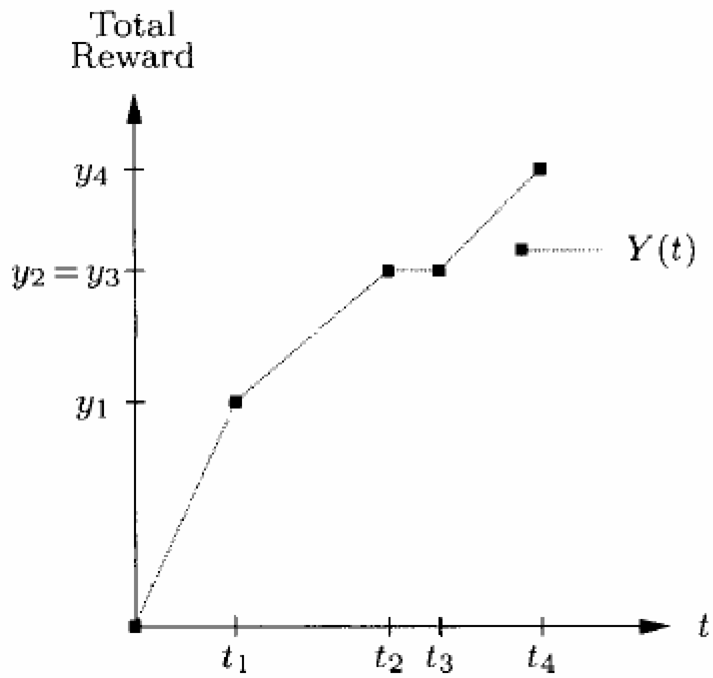
EXAMPLE



$r_0 = 3.0$ $r_1 = 1.0$ $r_2 = 0.0$







PERFORMABILITY

- Based on the definitions of *X(t), Z(t) and Y(t) which are non-independent* random variables, various measures can be defined.
- The most general measure is referred to as the *performability*:

$$\Psi(y, t) = P[Y(t) \leq y]$$

PERFORMABILITY

- Unfortunately, the performability is difficult to compute for unrestricted models and reward structures.
- Smaller models can be analyzed via double Laplace transform.
- The same mathematical difficulties arise if the distribution Φ(y, *t) of the time-average accumulated reward is to be computed:*

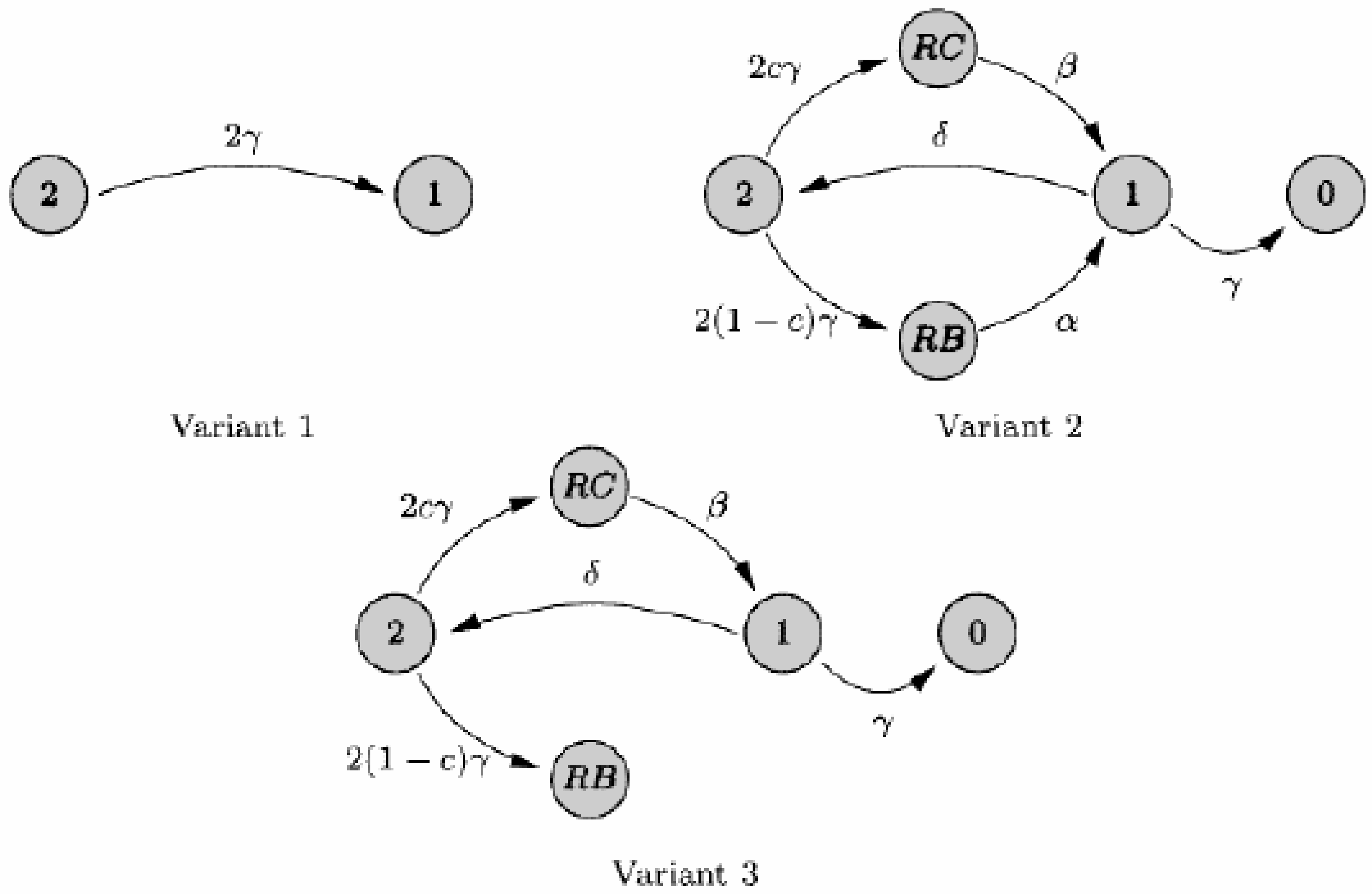
$$\Phi(y,t) = P\left[\frac{1}{t}Y(t) \le y\right]$$

EXPECTED REWARD RATE

- The *expected instantaneous reward rate can be computed from the following equation:*

$$E[Z(t)] = \sum_{i \in S} r_i \pi_i(t)$$

CASE STUDY

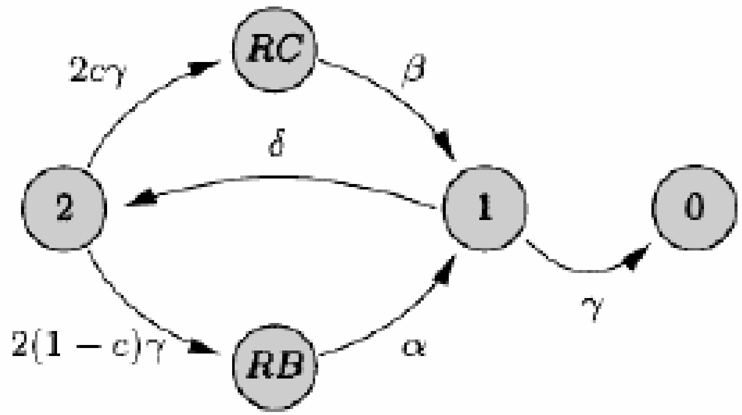


2
1
2γ
Variant 1
RC
RB
0
$2c\gamma$
β
δ
$2(1-c)\gamma$
α
γ
Variant 2
Variant 3

SYSTEM AVAILABILITY

- Availability measures are based on a *binary reward* structure.

State i	Reward Rate r_i
2	1
RC	0
RB	0
1	1
0	0



Variant 2

AVAILABILITY

- The instantaneous availability is given by:

$$A(t) = E[Z(t)] = \sum_{i \in S} r_i \pi_i(t) = \sum_{i \in U} \pi_i(t) = \pi_1(t) + \pi_2(t)$$

UNAVAILABILITY

- Unavailability can be calculated with a reverse reward assignment to that for availability.
- The instantaneous unavailability, therefore, is given by:

$$UA(t) = E[Z(t)] = \sum_{i \in S} r_i \pi_i(t) = \sum_{i \in D} \pi_i(t) = \pi_{RC}(t) + \pi_{RB}(t) + \pi_0(t)$$

CALL TO REPAIRS

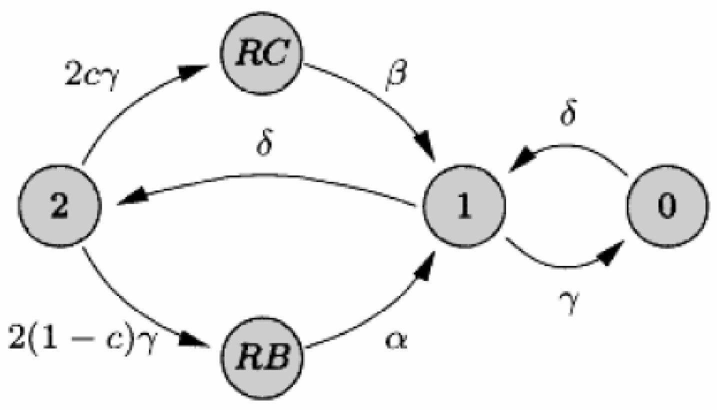
- Transient average number of repair calls N1(t)
- Steady state average number of repair calls N2(t)

$$N^{(1)}(t) = \frac{1}{t}E[Y(t)] = \delta(L_0(t) + L_1(t))$$

$$N^{(2)}(t) = E[Z] = t\delta(\pi_0 + \pi_1)\,.$$

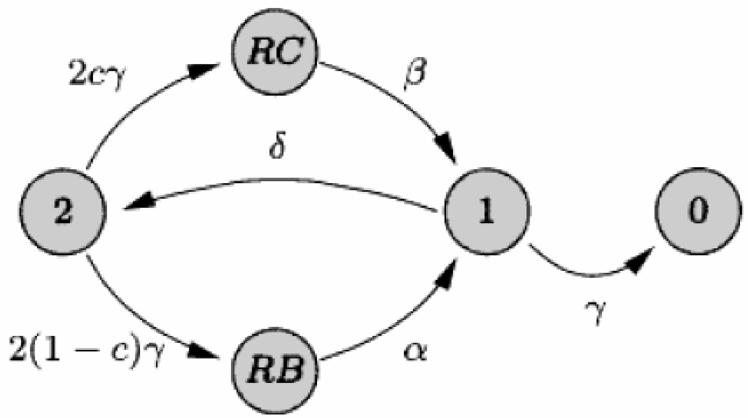
Table 2.3 Reward assignment for number of repair calls in $[0, t)$

State i	Reward Rate r_i
2	0
RC	0
RB	0
1	$t\delta$
0	$t\delta$



RELIABILITY

- Again, a *binary reward function r is defined that* assigns reward rates 1 to up states and reward rates 0 to (absorbing) down states, as given:



Variant 2

Table 2.5 Reward assignment for computation of reliability with variant-2 CTMC from Fig 2.7

State i	Reward Rate r_i
2	1
RC	1
RB	1
1	1
0	0

RELIABILITY

- Reliability can be expressed as:

$$R(t) = P[T > t] = P[Z(t) = 1] = 1 - P[Z(t) \leq 0] = E[Z(t)]$$

- Where the random variable *T characterizes the time to the next occurrence* of an unwanted (failure) event.

MTTA & MTTF

- With a known reliability function *R(t)*, *the mean time to the occurrence* of an unwanted (failure) event is given by

$$E[T] = \mathrm{MTTF} = \mathrm{MTTA} = \int_0^\infty R(t)\,\mathrm{d}\,t\,.$$

MTTF and MTTA are acronyms for *mean time to failure and mean time to absorption.*

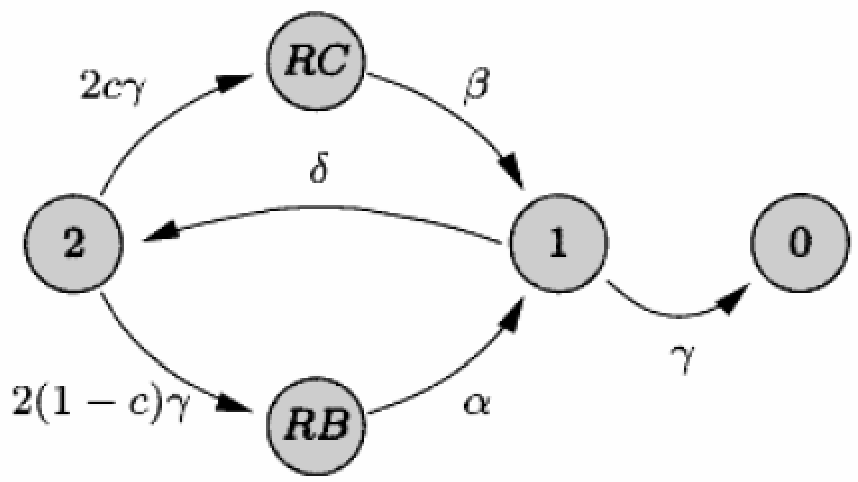
CATASTROPHIC EVENTS

- Related to reliability measures, we would often be interested in knowing the expected number of catastrophic events *C(t) to occur in a given time* interval [0, *t) . To this end*

$$C(t) = \frac{1}{t} E[Y(t)]$$

Table 2.6 Reward assignment for predicting the number of catastrophic incidents with variant-2 CTMC from Fig 2.7

State i	Reward Rate r_i
2	0
RC	0
RB	0
1	$t\gamma$
0	0

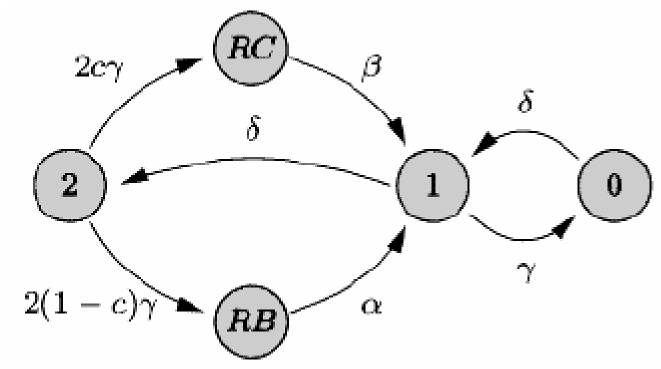


PERFORMANCE

- For computing availability and reliability most of the reward functions were binary ones.
- In this section, other examples would be considered.

TASK COMPLETION PROBABILITY

- Suppose there is a task which consumes x time units to be completed.
- We allow task requirement to be state dependent as well so that in state 2 the task requires x2 time units to complete and likewise for state 1



Task interruption probability:

$$P[O \leq x_2] = 1 - P[O > x_2] = 1 - e^{-2\gamma x_2}$$

TASK COMPLETION PROBABILITY

- Interruption probability IP(x1 , x2) of a task of length x1 , x2 , respectively is computed with E[Z]:

$$IP^{(1)}(x_1, x_2) = E[Z] = (1 - e^{-2\gamma x_2})\pi_2 + (1 - e^{-\gamma x_1})\pi_1$$

State i	Reward Rate r_i
2	$1 - e^{-2\gamma x_2}$
RC	0
RB	0
1	$1 - e^{-\gamma x_1}$
0	0

CONCLUSION

- MRMs as an extension to the Markov chains.
- Using MRMs to compute:
  - Availability
  - Reliability
  - Performance
- Reward Networks can be used as a high level modeling language to obtain MRMs.
- There exists also other high-level models to generate Reward Networks.

REFERENCES

- G. Bolch, S. Greiner, H. D. Meer and K. S. Trivedi, “Queuing Networks and Markov Chains”, John Wiley and Sons, Inc., Publication, 2006.