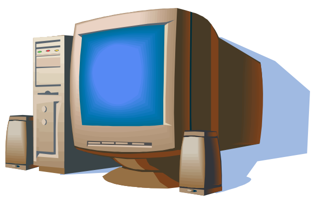
A Formalism for Stochastic Decision Processes with Asynchronous Events

Håkan L. S. Younes Reid G. Simmons
Carnegie Mellon University

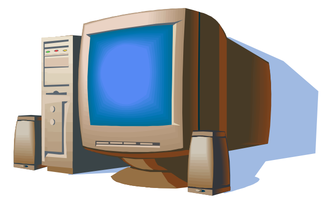
Introduction

- Asynchronous processes are abundant in the real world
  - Telephone system, computer network, etc.
- Discrete-time and semi-Markov models are inappropriate for systems with asynchronous events
- Generalized semi-Markov (decision) processes, GSM(D)Ps, are great for this!

Asynchronous Processes: Example



m_1

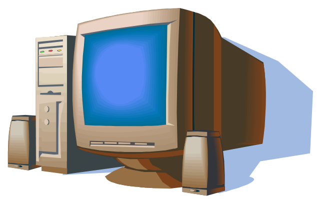


m_2

m_1 up
m_2 up

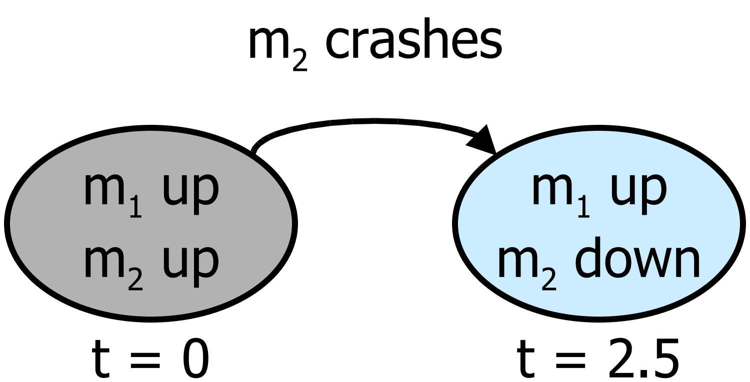
$t = 0$

Asynchronous Processes: Example

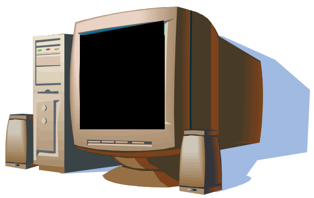


m_1

m_2

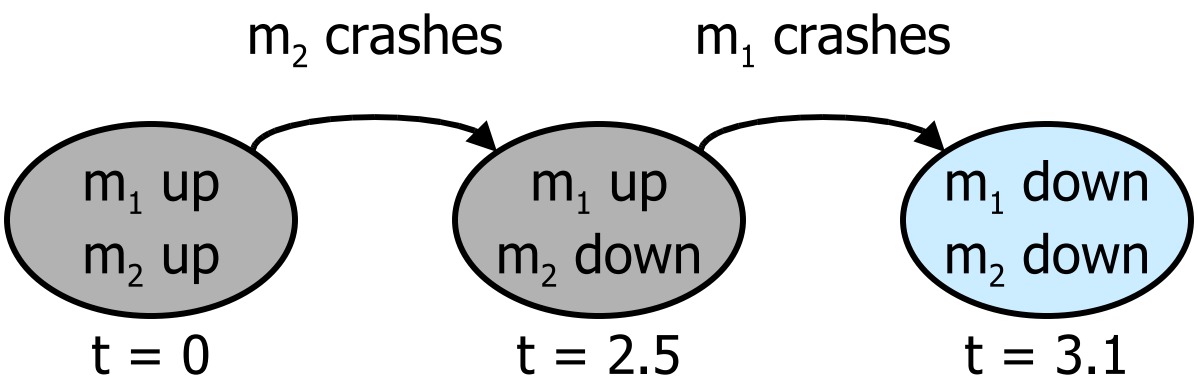


Asynchronous Processes: Example

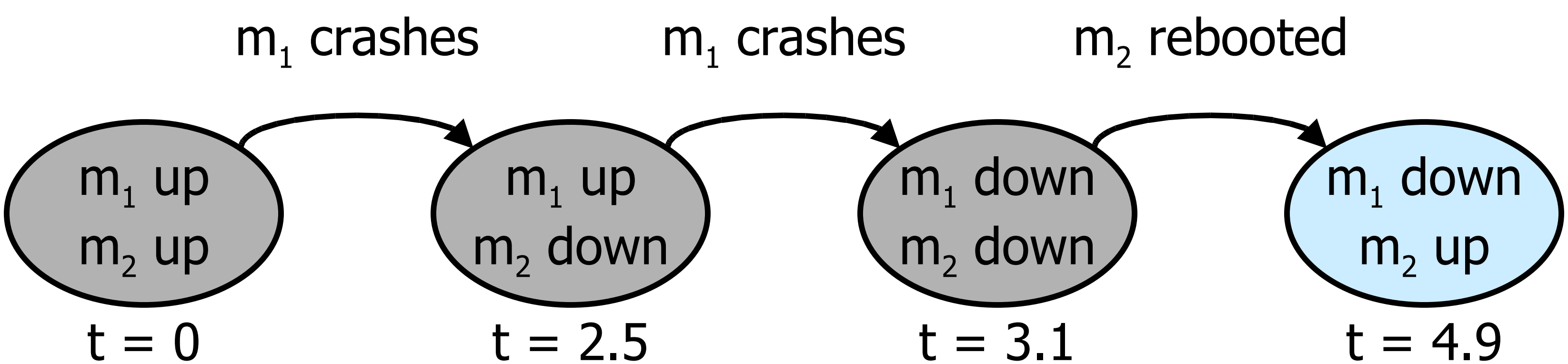


m_1

m_2



Asynchronous Processes: Example



A Model of Stochastic Discrete Event Systems

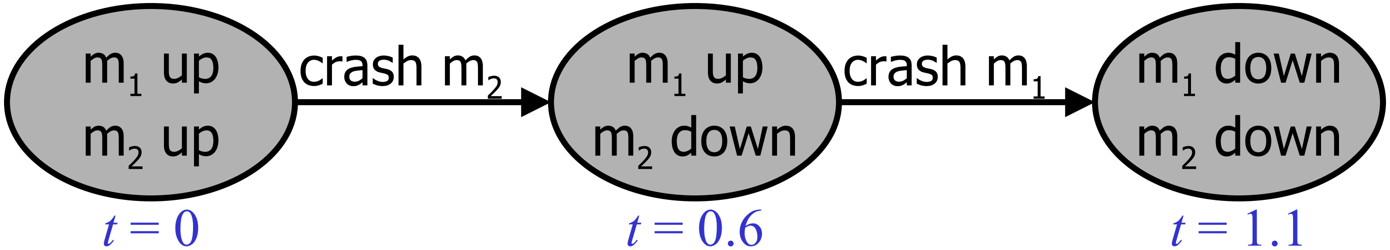
- Generalized semi-Markov process (GSMP) [Matthes 1962]
  - A set of events E
  - A set of states S
- GSMDP
  - Actions $A \subset E$ are controllable events

Events

- With each event e is associated:
  - A condition ϕ_e identifying the set of states in which e is enabled
  - A distribution G_e governing the time e must remain enabled before it triggers
  - A distribution $p_e(s'|s)$ determining the probability that the next state is s' if e triggers in state s

Events: Example

- Network with two machines
  - Crash time: $Exp(1)$
  - Reboot time: $U(0,1)$



$G_{c1} = Exp(1)$
$G_{c2} = Exp(1)$

$G_{c1} = Exp(1)$
$G_{r2} = U(0,1)$

$G_{r2} = U(0,0.5)$

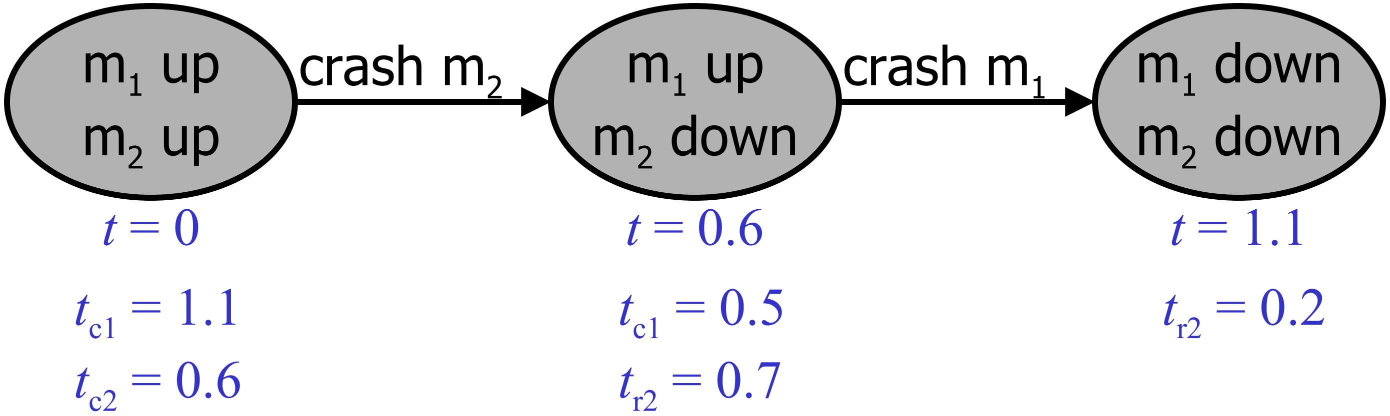
Asynchronous events $\Rightarrow$ beyond semi-Markov

Semantics of GSMP Model

- Associate a real-valued clock t_e with e
- For each $e \in E_s$ sample t_e from G_e
- Let $e^* = \text{argmin}_{e \in E_s} t_e$, $t^* = \min_{e \in E_s} t_e$
  - Sample s' from $p_{e^*}(s'|s)$
  - For each $e \in E_{s'}$
    - $t_e' = t_e - t^*$ if $e \in E_s \setminus \{e^*\}$
    - sample t_e' from G_e otherwise
- Repeat with $s = s'$ and $t_e = t_e'$

Semantics: Example

- Network with two machines
  - Crash time: $Exp(1)$
  - Reboot time: $U(0,1)$

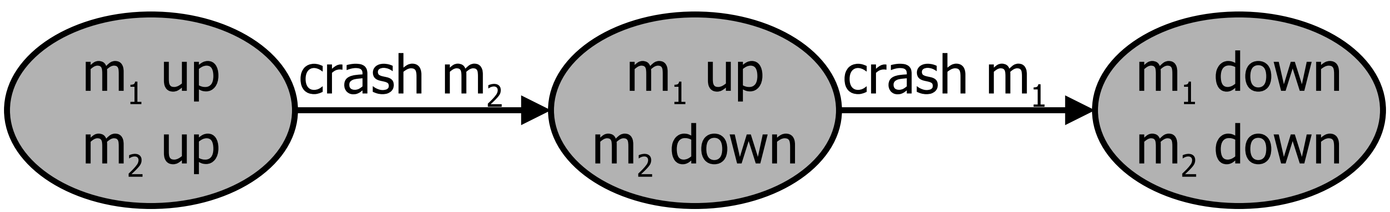


General State-Space Markov Chain (GSSMC)

- Model is Markovian if we include the clocks in description of states
  - Extended state space X
  - Next-state distribution $f(x'|x)$ well-defined
- Clock values are not know to observer
  - Time events have been enabled is know

Observations: Example

- Network with two machines
  - Crash time: $Exp(1)$
  - Reboot time: $U(0,1)$



$t = 0$ | $t = 0.6$ | $t = 1.1$

$t_{c1} = 1.1 \quad u_{c1} = 0$

$t_{c2} = 0.6 \quad u_{c2} = 0$

$t_{c1} = 0.5 \quad u_{c1} = 0.6$

$t_{r2} = 0.7 \quad u_{r2} = 0$

$t_{r2} = 0.2 \quad u_{r2} = 0.5$

Observation Model

- An observation o is a state s and a real value u_e for each event representing the time e has currently been enabled
- Summarizes execution history
  - $f(x|o)$ is well-defined
- Observations change continuously!

Policies

- Actions as controllable events
  - We can choose to disable an action even if its enabling condition is satisfied
- A policy determines the set of actions to keep enabled at any given time during execution

Rewards and Optimality

- Lump sum reward $k(s,e,s')$ associated with transition from s to s' caused by e
- Continuous reward rate $r(s,A')$ associated with $A' \subset A$ being enabled in s
- Infinite-horizon discounted reward
  - Unit reward earned at time t counts as $e^{-\alpha t}$
- Optimal choice may depend on entire execution history
  - Policy maps observations to sets of states

Shape of Optimal Policy

- Continuous-time MDP: $G_e = Exp(\lambda_e)$ for all events
  - Sufficient to change action choice at the time of state transitions
- General distributions
  - Action choice may have to change between state transitions (see [Chitgopekar 1969, Stone 1973, Cantaluppi 1984] for SMDPs)

Solving GSMDPs

- Optimal infinite-horizon GSMDP planning: feasible?
- Approximate solution techniques
  - Discretize time
  - Approximate general distributions with phase-type distributions and solve resulting MDP [Younes & Simmons, AAAI-04]

Summary

- Current research in decision theoretic planning aimed at synchronous systems
- Asynchronous systems have largely been ignored
- The generalized semi-Markov decision process fills the gap