Comparing the speed of probabilistic processes

Mathias R. Pedersen[1], Nathanaël Fijalkow[2], Giorgio Bacci[1], Kim G. Larsen[1], and Radu Mardare[1]

[1]Department of Computer Science, Aalborg University, Denmark

[2]Alan Turing Institute, London, United Kingdom

Semi-Markov processes

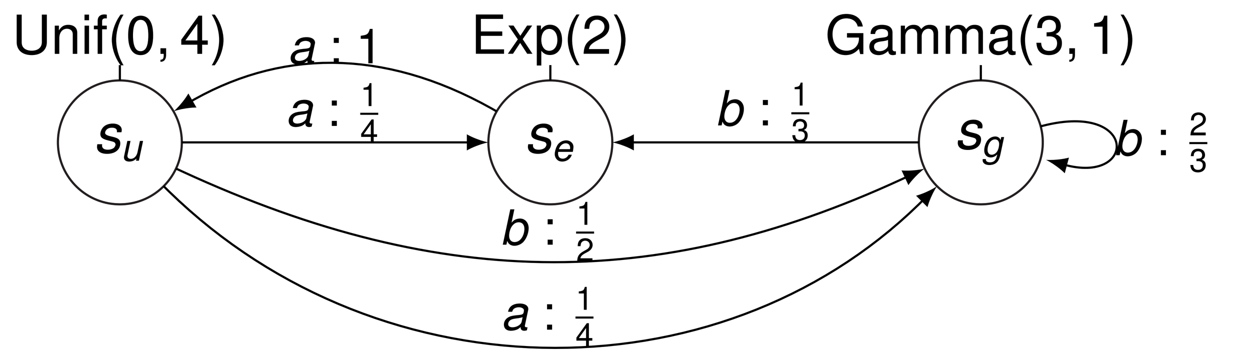

Fig. 1: A semi-Markov process with different residence-time distributions.

Semi-Markov processes (SMPs) are real-time stochastic processes in which the time that is spent in a given state before a transition is fired is determined by an arbitrary probability distribution.

- **Continuous-time Markov chains** are a special case of SMPs where all residence-time distributions are exponential.
- SMPs have been used extensively to model systems where the time behaviour is not exponentially distributed.

Faster-than relation

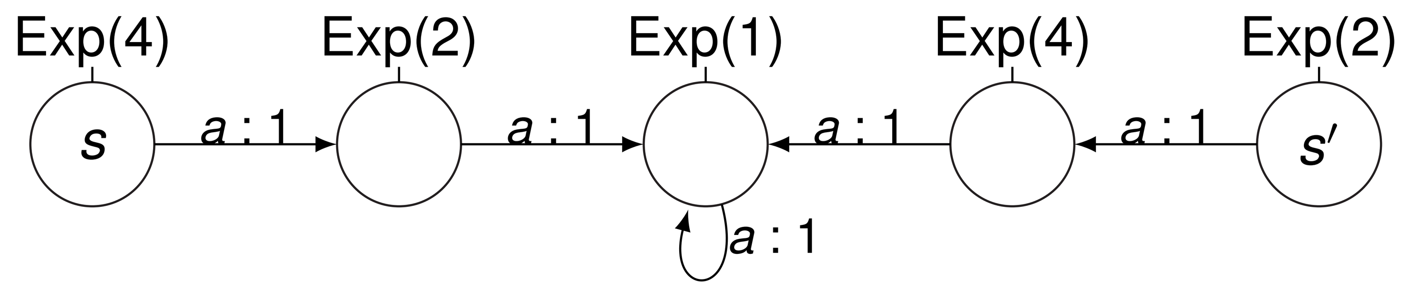

Fig. 2: s is faster than s'.

- Our goal is to **compare the speed** of processes.
- We will consider **trace-based** semantics of SMPs.

Definition 1. s is **faster than** s' if for any word w and time point t, s has a higher probability of outputting w within time t than s' does.

Example 1. Consider the SMP in Fig. 2. Let

$$X_1 \sim \text{Exp}(4), X_2 \sim \text{Exp}(2), X_3 \sim \text{Exp}(1)$$

be random variables. Denote by $\mathbb{P}_s(w, X \leq t)$ the **probability** that s outputs the word w and the random variable X is less than or equal t.

$$\mathbb{P}_s(a, X_1 \leq t) \geq \mathbb{P}_{s'}(a, X_2 \leq t)$$

because X_1 has a higher rate than X_2.

$$\mathbb{P}_s(aa, X_1 + X_2 \leq t) = \mathbb{P}_s(aa, X_2 + X_1 \leq t) = \mathbb{P}_{s'}(aa, X_2 + X_1 \leq t)$$

because addition of random variables is commutative, and hence also

$$\mathbb{P}_s(aaa^n, X_1 + X_2 + X_3^n \leq t) = \mathbb{P}_{s'}(aaa^n, X_2 + X_1 + X_3^n \leq t).$$

Therefore s is **faster than** s'.

Hardness results

Through a connection to the **Universality problem** for probabilistic automata, we obtain the following **undecidability result**.

Theorem 1. *The faster-than relation is undecidable.*

Since the Universality problem with only one letter is equivalent to the **Positivity problem** for linear recurrence sequences (which is related to the **Skolem problem**) [1], we also get a **hardness result**.

Theorem 2. *The faster-than relation is Positivity-hard for only one label.*

Furthermore, utilising a celebrated theorem for probabilistic automata by Condon and Lipton [2], we have an **inapproximability result**.

Theorem 3. *The faster-than relation can not be approximated up to a multiplicative constant.*

Unambiguous processes

A SMP is **unambiguous** if every output label leads to a unique successor state.

Example 2. The SMP in Fig. 1 is not unambiguous, since s_u has a-transitions to both s_e and s_g, whereas the SMP in Fig. 2 is unambiguous.

For unambiguous SMPs we can recover decidability.

Theorem 4. *For unambiguous SMPs, the faster-than relation is decidable in* **coNP**.

Given a state space S and states s and s', the **algorithm** is as follows:

- Using a **simple graph analysis**, find all the states p and p' reachable from s and s' such that there is a **looping** word w that takes p to p and p' to p'.
- Check whether s is faster than s' for all words of length less than $|S|^2$ and p is faster than p' for all looping words of length less than $|S|^2$.

Example 3. Consider the SMP in Fig. 3.

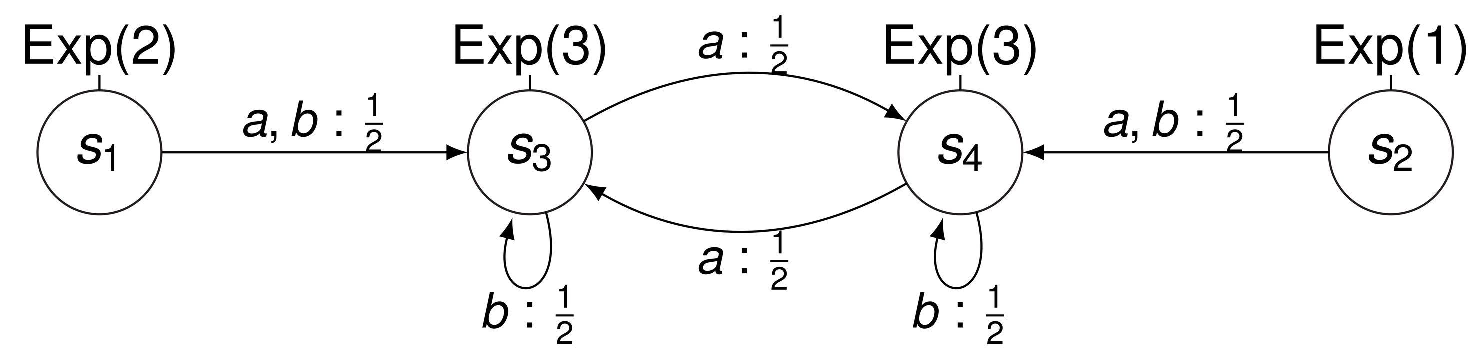

Fig. 3: An unambiguous semi-Markov process.

We want to check whether s_1 is faster than s_2.

- From (s_1, s_2) we can reach (s_3, s_4) which has looping words of the form $b^n ab^n a$, and we can also reach (s_4, s_3) with the same looping words.
- One can easily check that s_1 is faster than s_2 and that s_3 is faster than s_4 and vice versa for all looping words.
- However, we can also see that s_1 is faster than s_2 if and only if s_3 and s_4 have the same rate, for if s_3 had a higher rate, then the looping word aba would have a higher probability in s_4 than in s_3. Likewise, if s_4 had a higher rate, the looping word baa would fail the check.

Time-bounded approximation

Under the following assumptions, we can recover approximability:

- approximation up to an **additive** constant,
- consider only time up to a given **time bound**,
- **slow** distributions that must use some amount of time.

Under these assumptions we get the following:

- Slow distributions must use some non-zero amount of time in each step, so long words have a very small probability of being output within the given time bound.
- Hence words above a given length will have probability less than the desired approximation accuracy.
- We therefore need only check words up to a given length.

Theorem 5. *Approximating the time-bounded faster-than relation up to an additive constant is possible.*

Open problems

There are a number of open problems still:

- The symmetric **equally-fast** relation.
- **Reactive** models instead of **generative** models.
- **Logical** aspects of the relation.
- **Compositional** aspects of the relation, including **timing anomalies**.

References

[1] S. Akshay, Timos Antonopoulos, Joël Ouaknine, and James Worrell. Reachability problems for Markov chains. **Inf. Process. Lett.**, 115(2):155–158, 2015.

[2] Anne Condon and Richard J. Lipton. On the complexity of space bounded interactive proofs (extended abstract). In **FOCS**, 1989.