Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Computer Engineering and Networks Laboratory
Computer Engineering Group (TEC)

HS 2019 Prof. L. Thiele / X.He

# Discrete Event Systems

## Solution to Exercise Sheet 12

This last exercise session puts somehow everything together. You will discover a simulation tool for time Petri nets, which is paramount for the evaluation of your design. You will mainly use the "stepper simulator", which randomly plays the token game. In the last part of the exercise, you will use the embedded model-checker to tune some design parameters.

## 1 Warming up: Calculating with Petri nets

In this exercise you are supposed to model a function $f_i(x, y)$ using a Petri net. That is, the Petri net must contain two places $P_x$ and $P_y$ that hold $x$ and $y$ tokens respectively in the beginning. Additionally, the net must contain one place $P_z$ which holds $f_i(x, y)$ tokens when the net is dead. The Petri nets are supposed to work for arbitrary numbers of tokens in $P_x$ and $P_y$.

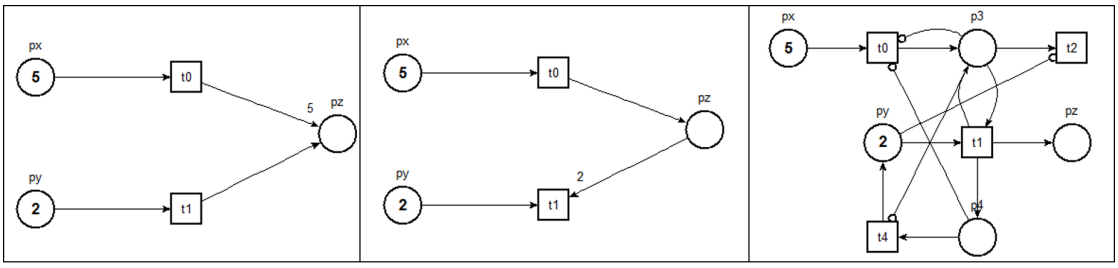
**a)** $f_1(x, y) = 5x + y \quad \forall x, y \geq 0$

**b)** $f_2(x, y) = x - 2y \quad \forall y \geq 0, x \geq 2y$

**c)** $f_3(x, y) = x \cdot y \quad \forall x, y \geq 0$

Here you may want to use inhibitor arcs. An inhibitor arc between a place and a transition prevents the transition from firing as long as there is at least one token in the place.
**Hint** Start by creating a net that "duplicate" the number of tokens from $P_x$ in place $P_z$. Then adapt this net to perform the multiplication.

There are multiple options here. I just present one set of solutions. These are captures of TINA models of these nets. The models are available with the solution.




a) b) c)

## 2 Simulate your Petri nets with TINA

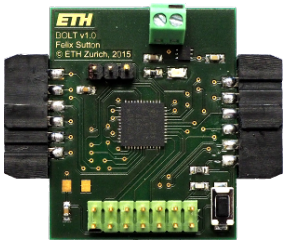
In this exercise, you will learn how to use the modeling tool TINA (**TI**me petri **N**et **A**nalyzer). The objective is to implement your solutions from Exercise 1 and simulate the nets.

**a)** On our course webpage you will find a simple tutorial for basic uses of the TINA software. Complete the reading of Section 1 before moving on.

**b)** To practice, create new net files and implement your solutions from Exercise 1. Using the stepper, try them out for different values of $x$ and $y$.

# 3 Queue sizing and overflow management of BOLT

BOLT is an ultra-low power processor interconnect that decouples arbitrary application and communication processors with respect to time, power and clock domains. BOLT supports asynchronous message passing with predictable timing characteristics. Therefore, it enables system designers to construct highly-customized platforms that are easier to design, implement, debug, and maintain. You can find more information on BOLT's webpage: www.bolt.ethz.ch



The BOLT processor interconnect

In this exercise, we look at the practical problem of queue sizing and overflow management, using BOLT architecture as a case study.

Assume BOLT is connected on one side to a sensor, which delivers data at a given frequency and writes into BOLT queue. At the other side, a communication processor is responsible for reading those messages from BOLT and sending them through a wireless sensor network. The problem is that the network is not always available (other processes also need the bandwidth) and the communication processors have limited memory to store messages before sending them. The BOLT queues can be used as temporary buffers.

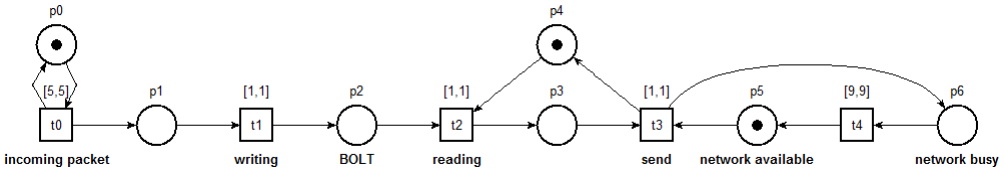
The objective is to evaluate the design parameters (queue size, resource management scheme...) such that overflows in the BOLT queue (which would result in packet losses) are avoided.

## 3.1 The nice and pretty deterministic world

**a)** This initial model (in the exercise folder) is untimed. Extend it with delays on transitions such that the net follows the earliest firing rule and knowing that:

- The sensor produces one message every 5 time units.
- Reading and writing from/to BOLT take one time unit each.
- Sending a message to the network takes one time unit.
- On average, the communication processor uses the network for 10% of the time.

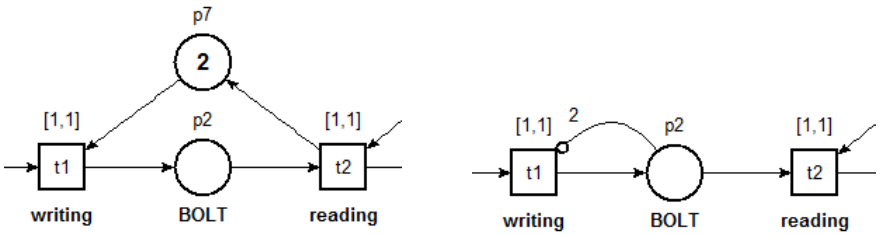
It should be rather easy to obtain the following model




**b)** Play the token game using the stepper. Is the net bounded? Why is this not surprising? This net is obviously not bounded because it generates one packet every 5 time units on one end, and disposes of one every 10 time units only on the other end. Hence, packets pile up in the middle.

**c)** Assume BOLT has now a maximal capacity of two messages. How would modify the net to implement this? Why is this not a "real" solution to our problem?

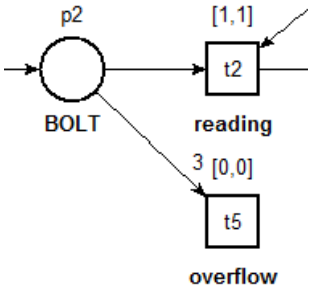
There are two easy ways of doing this. Either directly add a capacity to the `BOLT` place or use a weighted inhibitor arc from `BOLT` over the `writing` transition.




However, now BOLT is safe from overflow, but messages still pill up somewhere. In this case, "somewhere" is in place $p1$, which represents the sensor memory. Having this overflowing is not satisfactory either.

On the physical system, BOLT does not prevent a write operation if its queue is full, but overwrites old messages. Hence, our previous solution to avoid overflows is not satisfying, as it yields that the write operation is forbidden if there is no more memory available in BOLT, so the systems is actually waiting for the resource to be available.

In our case, the goal is to guarantee there is enough network resource available such that there is no possible overflow in BOLT, without having to control the write operations. As a result, we need our Petri net to model the (possible) overflow of BOLT in order to test our design. Among several options, we choose to do that by adding a sink transition downstream the `BOLT` place, as shown thereafter.



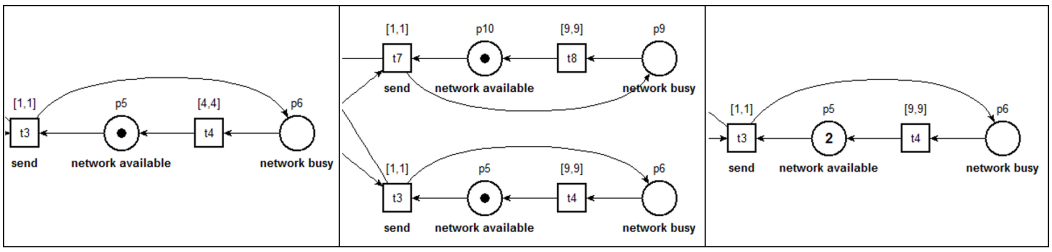

If it happens that 3 tokens are in the `BOLT` place, the overflow transition is enabled and can fire. This is used as a flag to detect an overflow of BOLT.

**e)** Add the `overflow` place to your model.

**f)** Let us try to avoid overflows by increasing the amount of `network` resource available. Say, we want to double it on average, such that it matches the generation of sensor data.. This can be done in several ways:

- Assume you get the network 20% of the time.
- Assume there are two networks and you get each for 10% of the time. Further assume the communication processor can read and store one message per network (e.g., if there are two networks, you should have 2 tokens in place $p4$). One can either
  - explicitly model two independent networks, or
  - set a second token in the `network available` place.

Those three options are illustrated below




a) b) c)

Try out the three options.

**Note 1.** You can have multiple nets open simultaneously. Just open multiple GUI in parallel.

**Note 2.** You can copy/paste parts of a net.

Why solution **c)** does not solve our problem? Which of solution **a)** or **b)** seems the more flexible to you? Why?

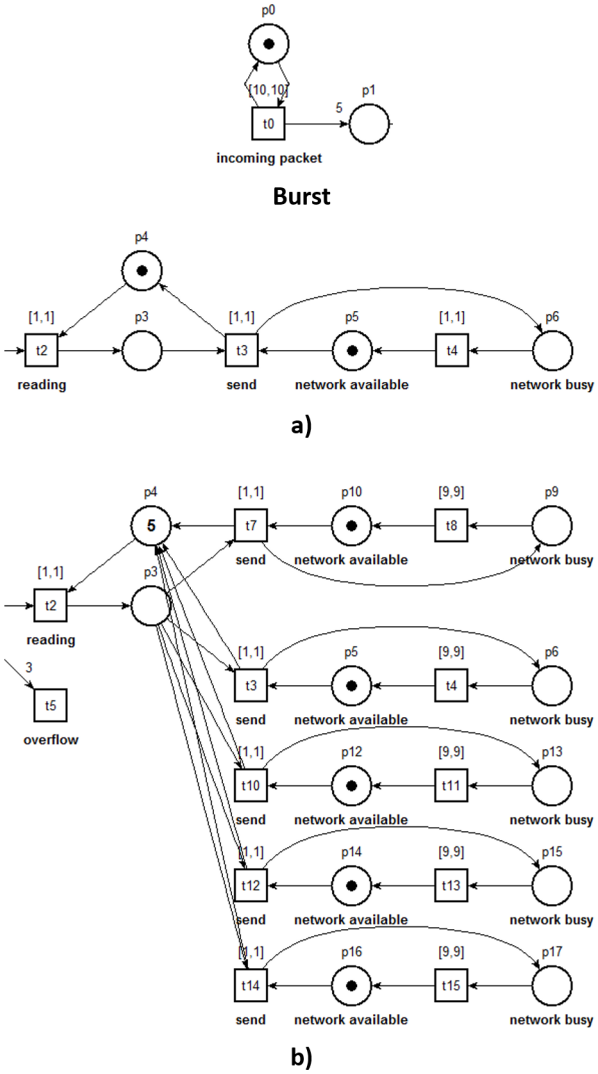
Option **c)** does not work because of the mechanism of delay reset in time Petri nets. When there are two tokens in `network busy` and `t4` fires, this creates an evolution of the enabling of `t4`, hence the delay is recomputed (i.e., the clock is reset, as explained in the lecture). Since the delay cannot be anything else than 9 time units, it starts over from there, regardless how long the second token has already spent in the place. This implies that the resource has in fact a cycle time of 9 time units, and not 5 as we wanted.

Furthermore, we will emphasis with the next question that solution **b)** is more flexible, as you can send 2 messages anytime during a 10 time units window, while for **a)**, you must wait 5 time units between two messages.

**g)** Assume now that incoming packets arrive in burst of 5 messages every 10 time units. Consider again solutions of type **a)** and **b)** with increased bandwidth:

- One network that you get 50% of the time.
- Five networks that you get 10% of the time each.

Model the bursts and try out these options. Which solution seems the more flexible to you now?




As mentioned before, option **b)** is more flexible, it can absorb bursts that **a)** cannot. The reason is that the communication processor cannot read out fast enough the messages from BOLT, one every two time units at the most. If you allow multiple reads (e.g., 5 tokens in p4), it makes options **a)** and **b)** equivalent.

As you can see, even in such basic scenario, the analysis of our system design is not so easy. Understanding of the impact of all design choices is hard. In practice, it is often impossible to take everything into account in one model for real-life systems. This is why test and verification methods have been developed and are common practice nowadays.

## 3.2 Real-world is non-deterministic

Let us now consider a slightly more realistic model. For each task, there is not one precise execution time anymore but a time interval. For example, for reading and writing from/to BOLT, this represents the Best- and Worst-Case Execution Time (BCET/WCET). The resource management mechanism and the communication processor are also different.

- The processor can read and store one or several messages from BOLT before sending them.
- When the processor accesses the network to send messages, it will keep on sending (i.e., no releasing the resource) until it has no more message left in memory.

- It is uncertain when the network will be available again.

**a)** Complete the reading of Section 2 of the TINA tutorial.

Our design question is the following:

**How much memory does one need in the communication processor in order to guarantee that BOLT never overflows?**

To answer this question, we will use the LTL model-checker embedded within TINA. LTL is a different logic than CTL, but for our simple purpose here, we can say the differences are the following

- the $E$ quantifier doesn't exist,
- the $A$ quantifier is always implied,
- $G$ writes $\Box$ , $F$ writes $\Diamond$, and $X$ writes $\circ$.

This is summarized in the table below:

| CTL | AF a | AG a | EF a | AX a | ... |
|---|---|---|---|---|---|
| LTL | $\Diamond$ a | $\Box$ a | $\neg$ AG $\neg$ a | $\circ$ a | ... |
| in TINA | `<> a ;` | `[] a ;` | `- [] - a ;` | `() a ;` | ... |

**b)** What does the LTL property $\Diamond t5$ mean in natural language? Does our model verify this property? Are we happy about this?

$\Diamond t5 \equiv$ "Whatever happens, eventually transition $t5$ will fire."
The initial model verifies this property. But this is the opposite of what we want!! It means there will always be an overflow of BOLT at some point, even in the best-case scenario...

**c)** What is the LTL property our system must verify in order to guarantee that BOLT will never overflow?

We want that "Whatever happens, there is no overflow of BOLT". This can be verified with the LTL query "$\Box - t5$" evaluating to TRUE. This is equivalent to the CTL "$AG\,not(t5)$".

**d)** What is the smallest capacity for the place $p4$ that makes this true? Do not open the stepper to generate an example trace, or be prepared to wait for a while...

**Hint** Look for a solution between 1 and 30.

For this model, the communication processor needs to be able to read at least 27 messages to absorb the worst-case of messages arrival in BOLT and of the network availability.