Concurrency Theory

Lecture 18: True Concurrency Semantics of Petri Nets (I)

Joost-Pieter Katoen and Thomas Noll

Lehrstuhl für Informatik 2
Software Modeling and Verification Group

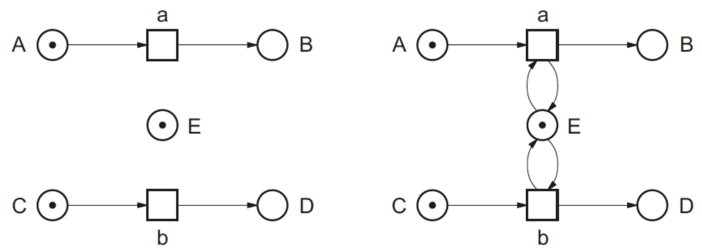
http://moves.rwth-aachen.de/teaching/ws-1718/ct

December 14, 2017

Overview

Overview

Motivation



Nets with identical sequential runs (*a* occurs before *b*, or vice versa), but the left net allows the simultaneous execution of *a* and *b* whereas the right one does not.

Interleaving semantics cannot distinguish these nets!

This requires a finer perspective on transition execution.

Overview

Nets

Net

A Petri net N is a triple (P, T, F) where:

- P is the countable set of places
- T is the countable set of transitions with $P \cap T = \varnothing$
- $F \subseteq (P \times T) \cup (T \times P)$ are the arcs.

Places and transitions are generically called nodes.
We assume that ${}^\bullet t$ and $t^\bullet$ are finite, for each $t \in T$.

Note that the set of places and transitions is countable, not necessarily finite (anymore).

Marking

A marking M of a net $N = (P, T, F)$ is a mapping $M : P \to \mathbb{N}$.
For net $N = (P, T, F)$ and marking M_0, the tuple (P, T, F, M_0) is called an elementary system net. M_0 is the initial marking of N.

Transition occurrence

Enabling and occurrence of a transition

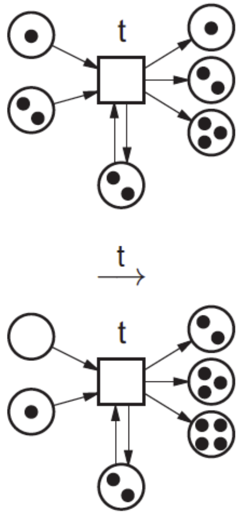
A marking M enables a transition t if $M(p) \geqslant 1$ for each place $p \in {}^\bullet t$.

Transition t can occur in marking M if t is enabled at M. Its occurrence leads to marking M', denoted $M \xrightarrow{t} M'$, defined for place $p \in P$ by:

$$M'(p) \;=\; M(p) - F(p,t) + F(t,p).$$

where we represent F by its characteristic function.

$M \xrightarrow{t} M'$ is also called a step of the net N.



Reachable markings

Step sequence

A sequence of transitions $\sigma = t_1\, t_2 \ldots t_n$ is a step sequence if there exist markings M_1 through M_n such that:

$$M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_2} \cdots \xrightarrow{t_{n-1}} M_{n-1} \xrightarrow{t_n} M_n$$

Marking M_n is reached by the occurrence of σ, denoted $M_0 \xrightarrow{\sigma} M_n$.

M is a reachable marking if there exists a step sequence σ with $M_0 \xrightarrow{\sigma} M$.

Sequential runs

Sequential run

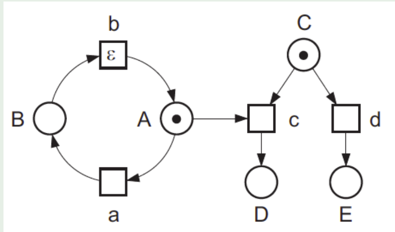
Let N be an elementary net system. A sequential run of N is a sequence

$$M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_2} \cdots$$

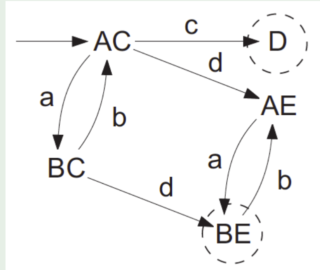
of steps of N starting with the initial marking M_0. A run can be finite or infinite. A finite run $M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_1} \cdot \xrightarrow{t_n} M_n$ is complete if M_n does not enable any transition.

Marking graph

The marking graph of N has as nodes the reachable markings of N and as edges the reachable steps of N.



A sample elementary net system



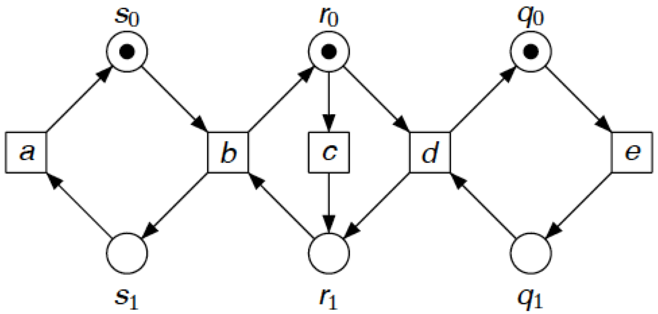
Its marking graph

The interleaving semantics of Petri nets

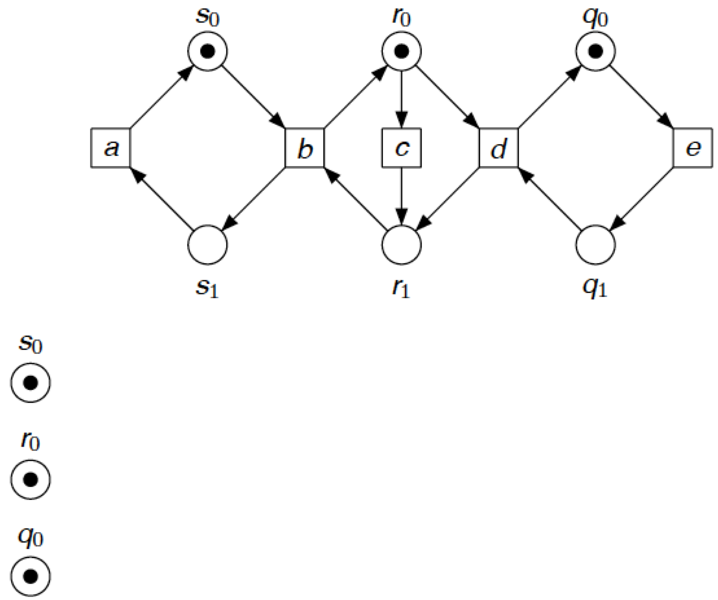
The interleaving semantics of a Petri net is its marking graph.

Overview

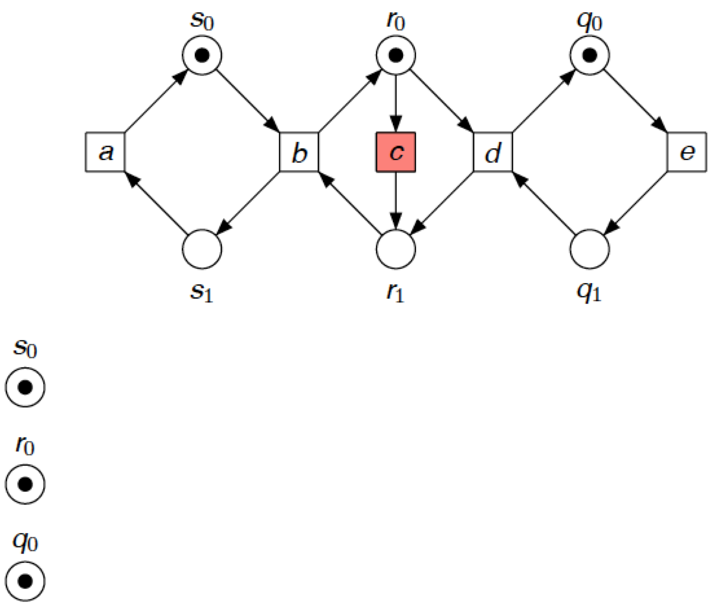
The true concurrency semantics of Petri nets



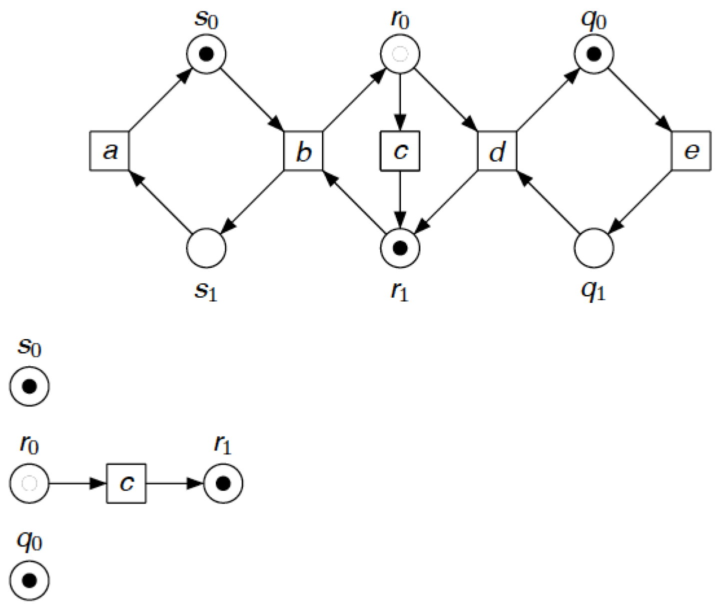
The true concurrency semantics of Petri nets



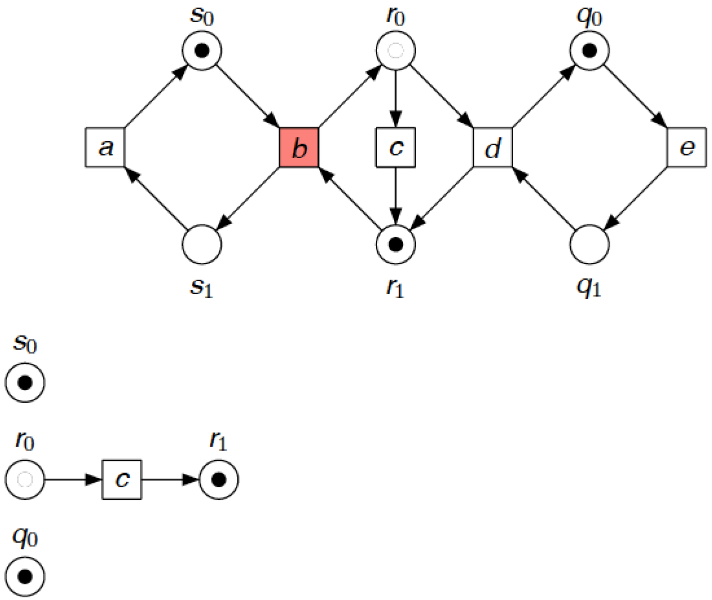
The true concurrency semantics of Petri nets



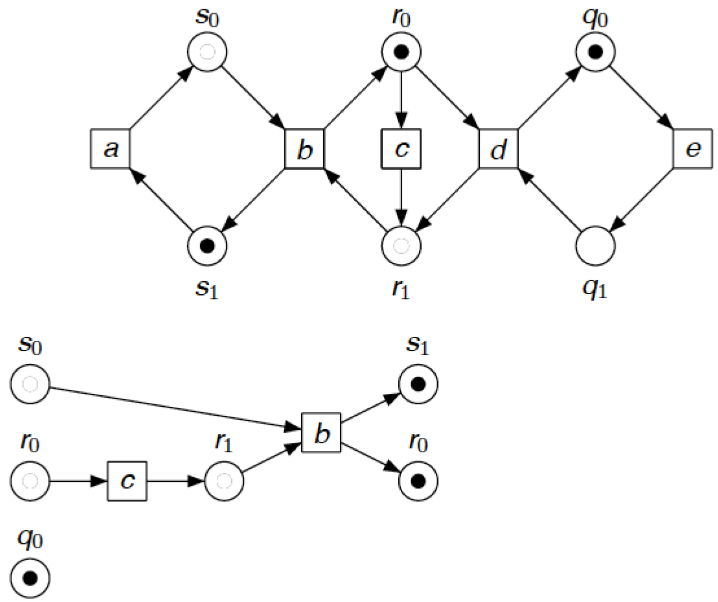
The true concurrency semantics of Petri nets



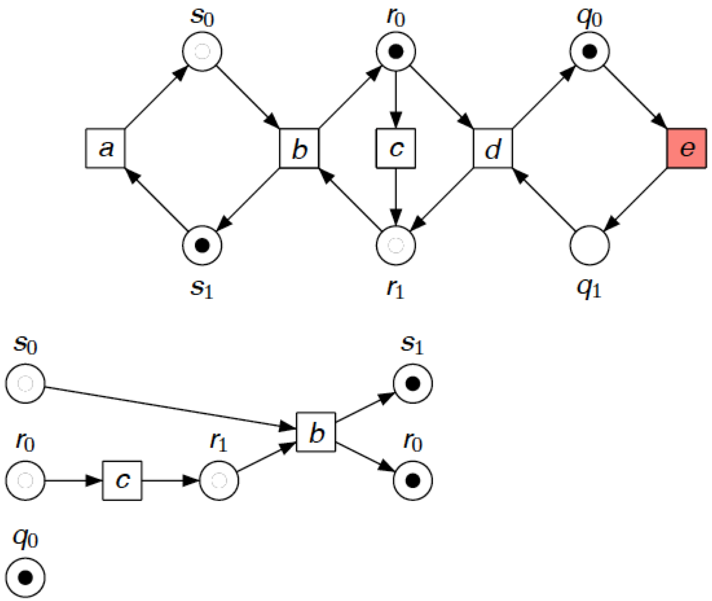
The true concurrency semantics of Petri nets



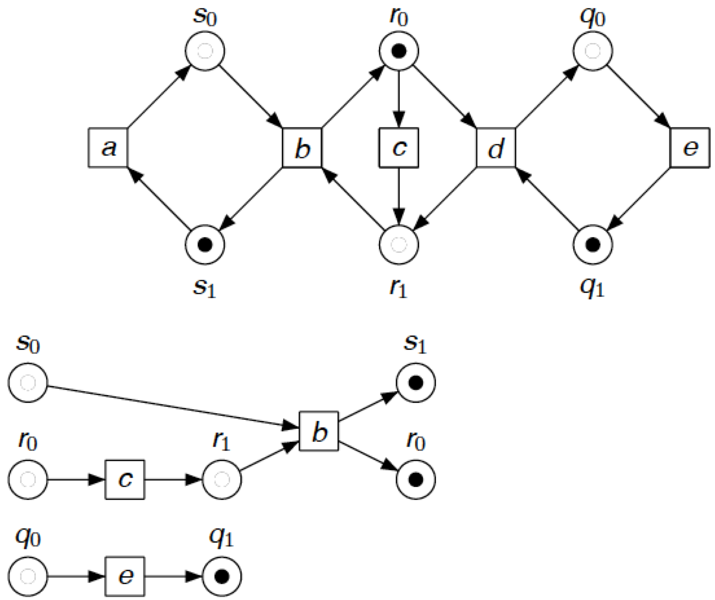
The true concurrency semantics of Petri nets



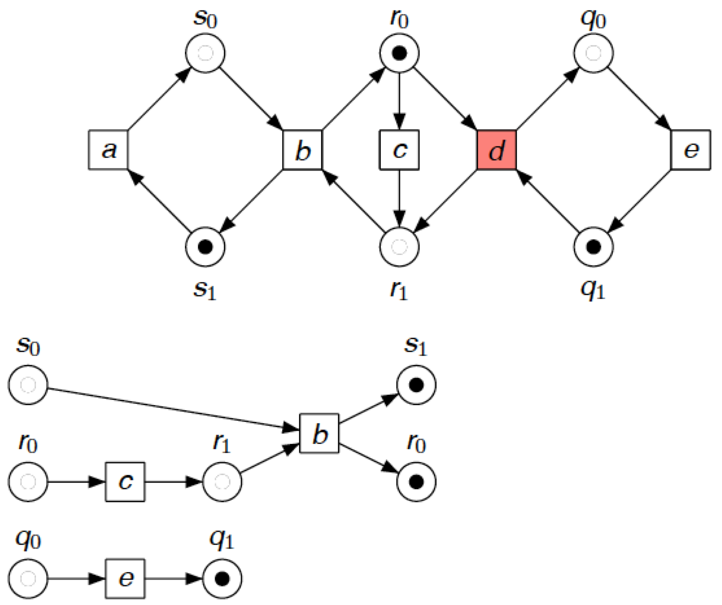
The true concurrency semantics of Petri nets



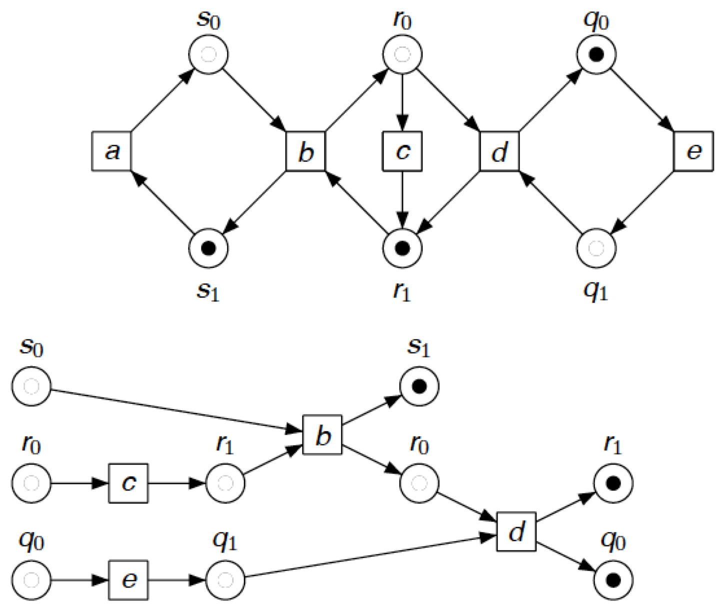
The true concurrency semantics of Petri nets



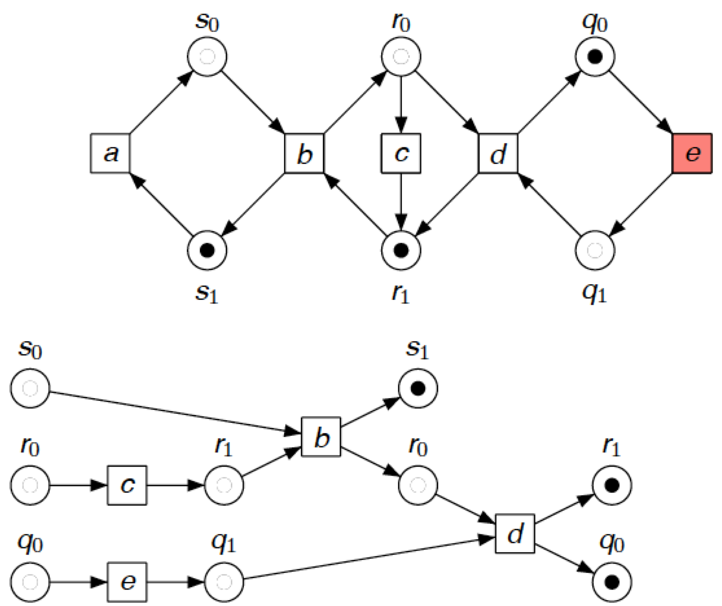
The true concurrency semantics of Petri nets



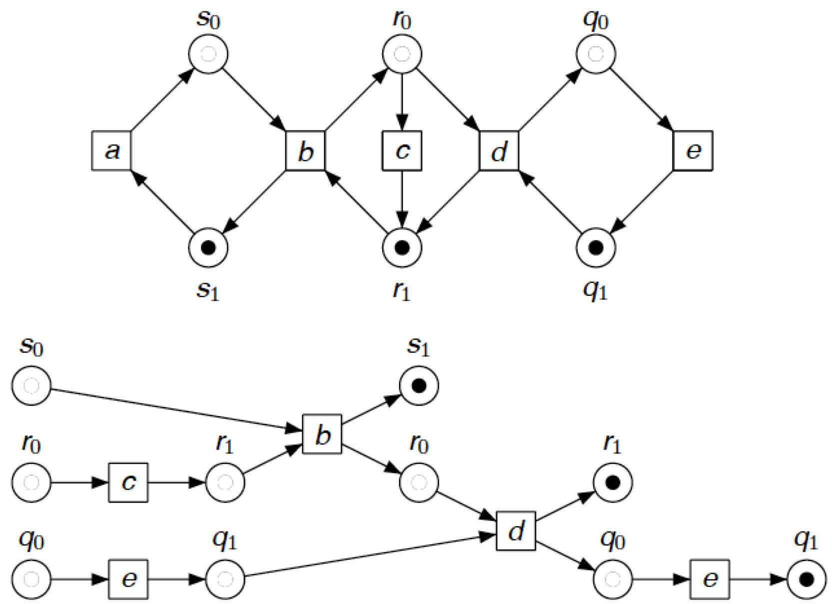
The true concurrency semantics of Petri nets



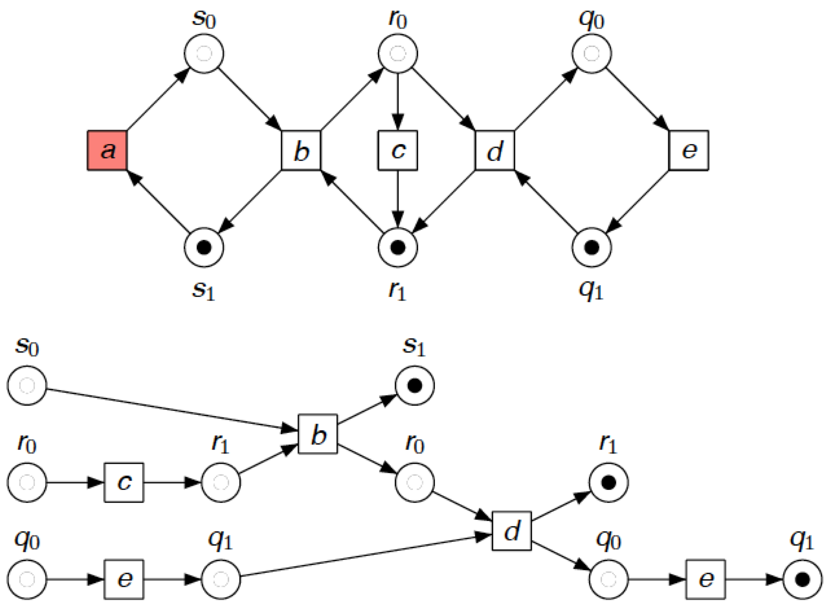
The true concurrency semantics of Petri nets



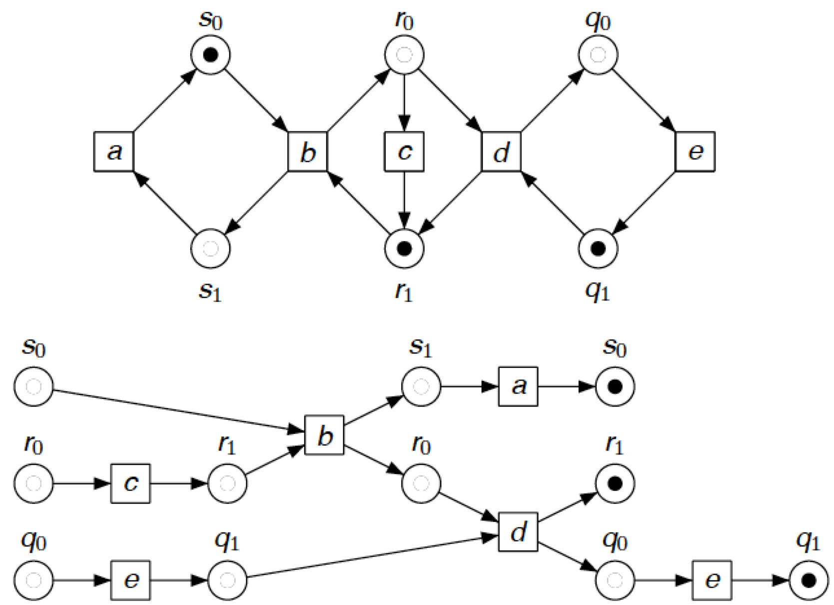
The true concurrency semantics of Petri nets



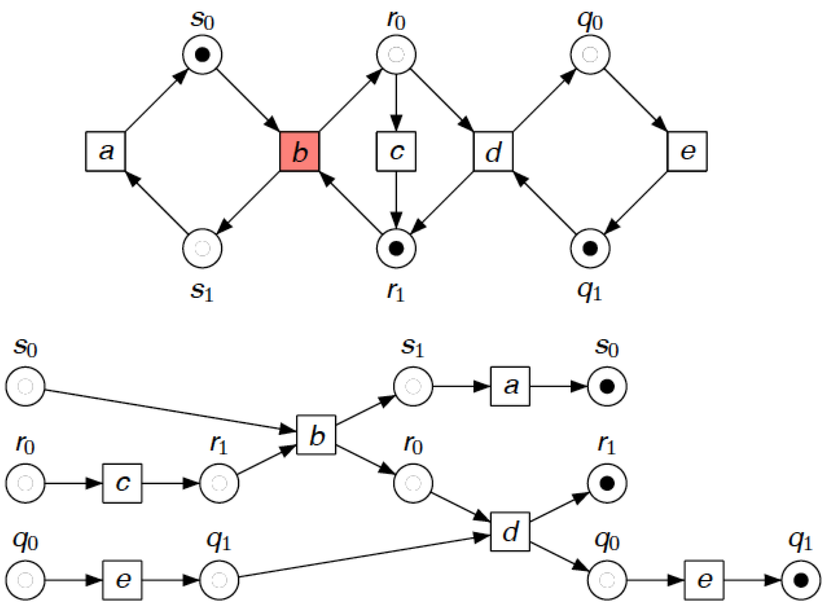
The true concurrency semantics of Petri nets



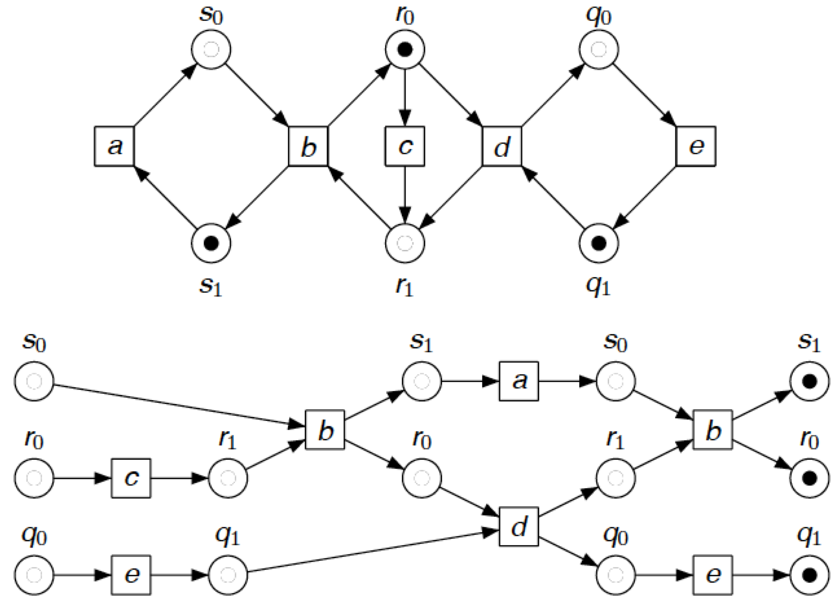
The true concurrency semantics of Petri nets



The true concurrency semantics of Petri nets



The true concurrency semantics of Petri nets



Interleaving versus true concurrency

The interleaving thesis:

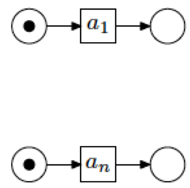
> *The total order assumption is a reasonable abstraction, adequate for practical purposes, and leading to nice mathematics*

The true concurrency thesis:

> *The total order assumption does not correspond to physical reality and leads to awkward representations of simple phenomena*

Interleaving versus true concurrency

In interleaving semantics, a system composed of n independent components



has $n!$ different executions

The automaton accepting them has 2^n states

In true concurrency semantics, it has only one nonsequential execution

Overview

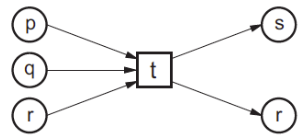
Actions

A distributed run of a net is a partial-order represented as a net whose basic building blocks are actions[1], simple nets

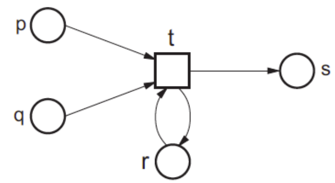
Action

An action is a labeled net $A = (Q, \{ v \}, G)$ with ${}^\bullet v \cap v^\bullet = \varnothing$ and ${}^\bullet v \cup v^\bullet = Q$.

Actions are used to represent transition occurrences of elementary net systems. If A represents transition t, then elements of Q are labeled with in- and output places of t and v is labeled t.

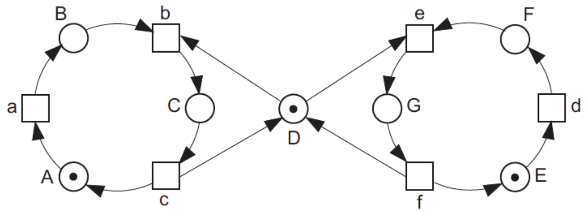


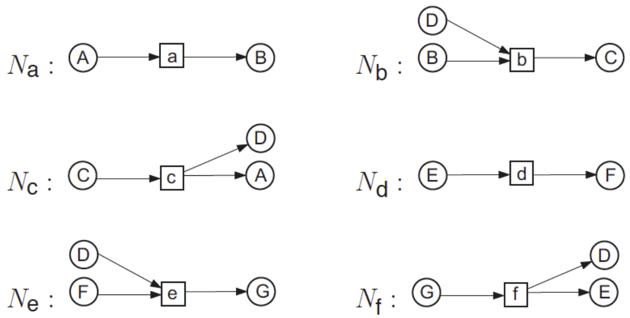
represents



[1]Not to be confused with the notion of action in transition systems.

Mutual exclusion net and its actions





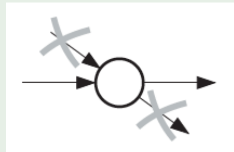
Causal nets

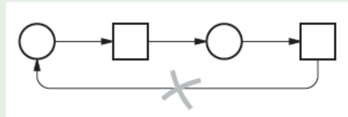
A causal net constitutes the basis for a "distributed" run.
It is a (possibly infinite) net which satisfies:

1. It has no place branches: at most one arc ends or starts in a place
2. It is acyclic
3. Each sequence of arcs (flows) has a unique first element
4. The initial marking contains all places without incoming arcs.

Intuition

No place branches, no sequence of arcs forms a loop, and each sequence of arcs has a first node.





Causal nets

A causal net constitutes the basis for a distributed run.
It is a possibly infinite net which satisfies:

1. Has no place branches: at most one arc ends or starts in a place
2. Is acyclic
3. Each sequence of arcs (flows) has a first element
4. The initial marking contains all places without incoming arcs

Causal net

A (possibly infinite) net $K = (Q, V, G, M_0)$ is called a causal net iff:

1. for each $q \in Q$, $|{}^\bullet q| \leqslant 1$ and $|q^\bullet| \leqslant 1$
2. the transitive closure (called causal order) G^+ of G is irreflexive
3. for each node $x \in Q \cup V$, the set $\{\, y \mid (y, x) \in G^+ \,\}$ is finite
4. M_0 equals the minimal set of places in K under G^+, i.e.,

$$M_0 \;=\; {}^\circ K \;=\; \{\, q \in Q \mid {}^\bullet q = \varnothing \,\}.$$

Note: the "runs" of the example net (with initial marking) are all causal nets.

Properties of causal nets

Lemma

Let $N = (P, T, F, M_0)$ be a causal net. Then every step sequence:

$$M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_2} \ldots\ldots \xrightarrow{t_k} M_k$$

of net N satisfies $M_j \cap t_k{}^\bullet = \varnothing$ for all $j = 0, \ldots, k{-}1$.

Proof.

By contraposition. Consider a step sequence of net N and suppose that $p \in M_j \cap t_k{}^\bullet$ for some $p \in P$ and some $0 \leqslant j < k$. This is impossible for $j = 0$, as by definition of causal nets, M_0 has no ingoing arcs, and thus $M_0 \cap t^\bullet = \varnothing$ for each $t \in T$. Hence, $j > 0$. Given that $p \in M_j$ (for some j) and $p \notin M_0$, it follows $p \in t_i{}^\bullet$ for some $0 < i \leqslant j$. (Some transition before reaching M_k must have put a token on p.) Thus $t_i, t_k \in {}^\bullet p$, where $t_i \neq t_k$ as F is well-founded. But by definition every place in a causal net is non-branching. So also p. Contradicting $t_i, t_k \in {}^\bullet p$ for $t_i \neq t_k$. □

Boundedness of causal nets

Lemma

Let $N = (P, T, F, M_0)$ be a causal net. Then every step sequence:

$$M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_2} \cdots\cdots \xrightarrow{t_k} M_k$$

of net N satisfies $M_j \cap t_k{}^\bullet = \varnothing$ for all $j = 0, \ldots, k{-}1$.

Boundedness of causal nets

Every causal net is one-bounded, i.e., in every marking every place will hold at most one token.

Proof.

Follows directly from the fact that the initial marking M_0 is one-bounded, and by the above lemma. □

Completeness of a causal net

Absence of superfluous places and transitions

Let $N = (P, T, F, M_0)$ be a causal net. Then there exists a possibly infinite step sequence

$$M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_2} \cdots\cdots \xrightarrow{t_k} M_k \xrightarrow{t_{k+1}} \cdots$$

of N such that $P = \bigcup_{k \geqslant 0} M_k$ and $T = \{ t_k \mid k > 0 \}$.

Proof.

On the black board. □

A causal net thus contains no superfluous places and transitions, as every place is visited and every transition is fired in the above step sequence.

Outset and end of a causal net

Outset and end of a causal net

The outset and end of causal net $K = (Q, V, G, M)$ are defined by:

$${}^\circ K = \{ q \in Q \mid {}^\bullet q = \varnothing \} \quad \text{and} \quad K^\circ = \{ q \in Q \mid q^\bullet = \varnothing \}.$$

Places without an incoming arc form the outset ${}^\circ K$. The places without an outgoing arc form the end K°.

What is a distributed run?

Distributed run

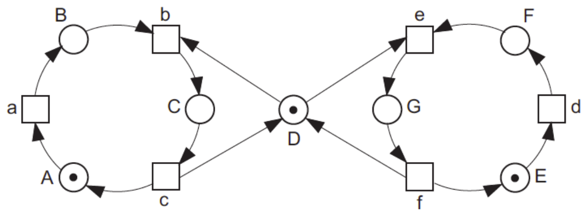
A distributed run of a one-bounded elementary net system N is:

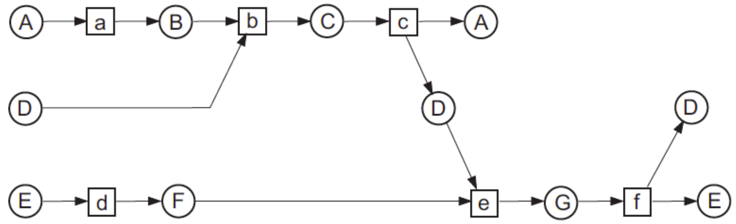
1. a labeled causal net K_N
2. in which each transition t (with ${}^\bullet t$ and $t^\bullet$) is an action of N.

A distributed run K_N of N is complete iff (the marking) ${}^\circ K$ represents the initial marking of N and (the marking) K_N° does not enable any transition.

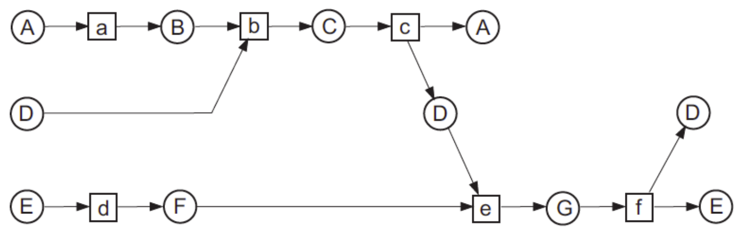
If N is clear from the context we just write K for K_N.

A distributed run for mutual exclusion





A distributed run for mutual exclusion



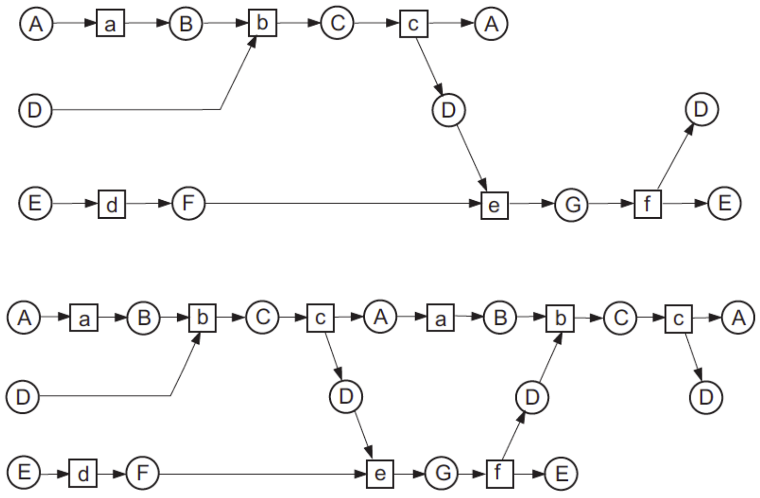
Distributed run of the mutual exclusion algorithm.

Actions N_a, N_b, N_c and N_d causally precede N_e. They form a chain.

N_a and N_d are not linked by actions; they are causally independent.

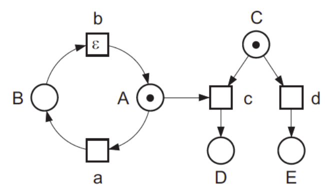
The same applies to N_b and N_d and N_c and N_d.

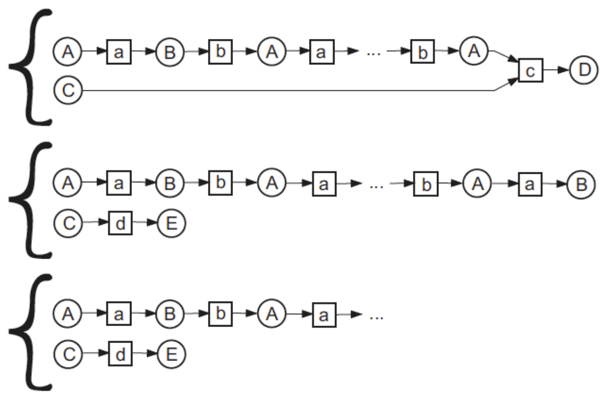
Expansion of a distributed run for mutual exclusion



A distributed run (top) and its extension with actions b and c.

More distributed runs

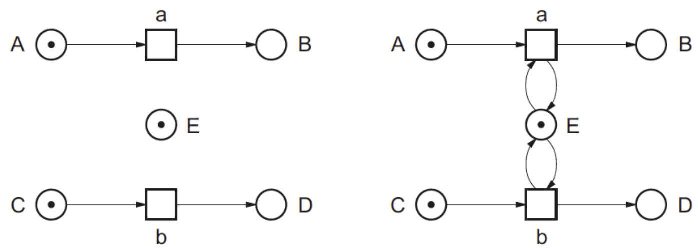




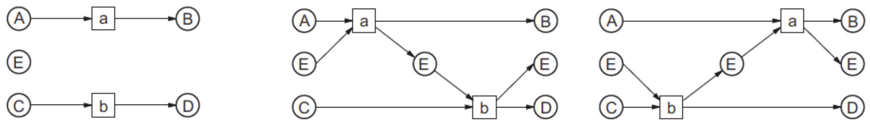
Various finite distributed runs and an infinite distributed run (right) of net (left).

Causal order

Opposed to sequential runs, distributed runs show the causal order of actions.



Nets with identical sequential runs (a occurs before b, or vice versa), but the left net has the left distributed run below, the right net both other ones:



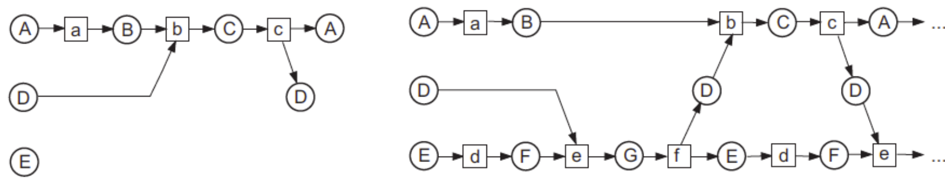
Composition of distributed runs

Composition of distributed runs

For $i = 1, 2$, let $K_i = (Q_i, V_i, G_i)$ be causal nets, labeled with ℓ_i. Let $(Q_1 \cup V_1) \cap (Q_2 \cap V_2) = K_1^\circ = {}^\circ K_2$ and for each place $p \in K_1^\circ$ let $\ell_1(p) = \ell_2(p)$. Then the composition of K_1 and K_2, denoted $K_1 \bullet K_2$, is the causal net $(Q_1 \cup Q_2, V_1 \cup V_2, G_1 \cup G_2)$ labeled with ℓ with $\ell(x) = \ell_i(x)$.

Intuition

The composition $K \bullet L$ is formed by identifying the end K° of K with the outset ${}^\circ L$ of L. To do this, K° and ${}^\circ L$ must represent the same marking.



What is a distributed run?

Distributed run

A distributed run of a one-bounded elementary net system N is:

1. a labeled causal net K
2. in which each transition t (with ${}^\bullet t$ and $t^\bullet$) is an action of N.

A distributed run K of N is complete iff (the marking) ${}^\circ K$ represents the initial marking of N and (the marking) K° does not enable any transition.

Examples on the black board.

Today: a characterization of distributed runs using homomorphisms.

Net homomorphisms

Homomorphism

A homomorphism from $N_1 = (P_1, T_1, F_1, M_{0,1})$ to $N_2 = (P_2, T_2, F_2, M_{0,2})$ is a mapping $h : P_1 \cup T_1 \to P_2 \cup T_2$ such that: [2]

1. $h(P_1) \subseteq P_2$ and $h(T_1) \subseteq T_2$, and
2. $\forall t \in T_1$, the restriction of h to ${}^\bullet t$ is a bijection between ${}^\bullet t$ (in N_1) and ${}^\bullet h(t)$ (in N_2), and similarly for $t^\bullet$ and $h(t)^\bullet$, and
3. the restriction of h to $M_{0,1}$ is a bijection between $M_{0,1}$ and $M_{0,2}$.[3]

Intuition

A homomorphism is a mapping between nets that preserves the nature of nodes and the environment of nodes. A homomorphism from N_1 to N_2 means that N_1 can be folded onto a part of N_2, or in other words, that N_1 can be obtained by partially unfolding a part of N_2.

[2] Here $h(X)$ for set X of nodes is defined by $h(X) = \bigcup_{x \in X} h(x)$.

[3] Due to the 1-boundedness, a marking M is a subset of the set P of places.

Distributed run using homomorphisms

Distributed run **[Best and Fernandez, 1988]**

A distributed run of an elementary net system N is a pair (K, h) where K is a causal net and h is a homomorphism from K to N.[4]

Intuition

A distributed run (K, h) of N may be viewed as a net K of which the places and transitions are labeled by places and transitions of N such that the labeling h forms a net homomorphism from K to N.[5]

[4] Best and Fernandez called this a process of a net.
[5] In the previous lecture, the labeling h was explicitly given as ℓ.

Examples

Overview

Summary

- A causal net is a possibly infinite net which is:
  - well-founded, acyclic, and has no place branching, and
  - whose initial marking are the places without incoming arcs
- Causal nets are one-bounded, and contain no redundant nodes
- A distributed run of N is a causal net whose nodes are labeled with nodes from N
- A distributed run can be obtained by composing causal nets
- Nets that have the same causal nets are causally equivalent
- Distributed run = the "true concurrency" analogue to a sequential run