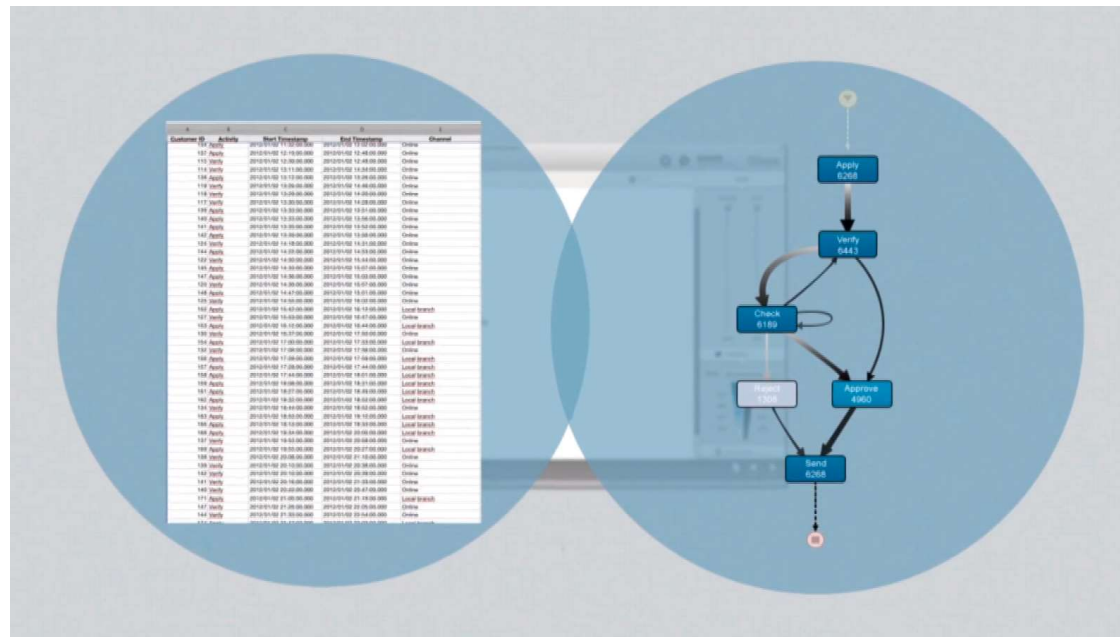


Folyamatbányászat 2

Esemény logok és folyamati modellek,
Petri hálók, Petri háló tulajdonságai
Átmenet rendszerek



Case	Activity	Timestamp	Resource
432	register travel request (a)	18-3-2014:8.16	John
432	get support from local manager (b)	18-3-2014:8.26	Mary
432	check budget by finance (d)	18-3-2014:8.66	John
432	decide (e)	18-3-2014:8.98	Sue
432	accept request (g)	18-3-2014:8.48	Mary

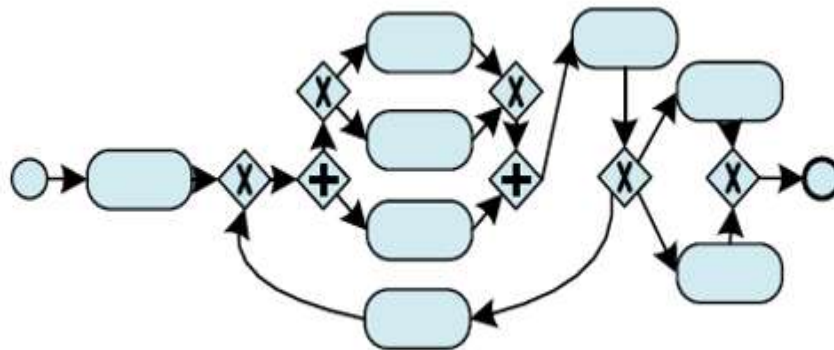


Valós lefutások modelljei és esemény logok

Diagram by W. van der Aalst

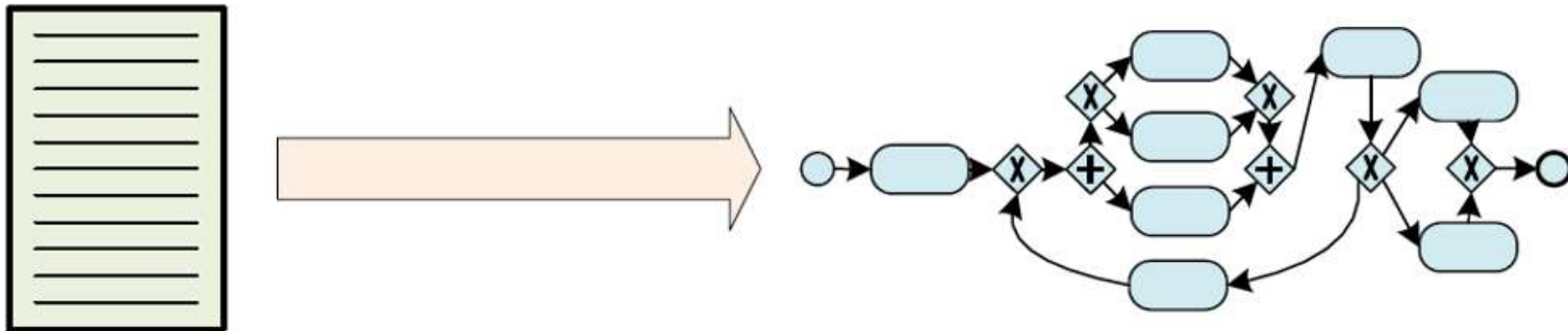
Play-Out

- Szimuláció
- Folyamat automatizálás
- Menedzsment játékok
- Modell ellenőrzés




Play-In

Folyamat felfedezés: a megfigyelt viselkedés alapján a folyamati modell előállítás, felépítése

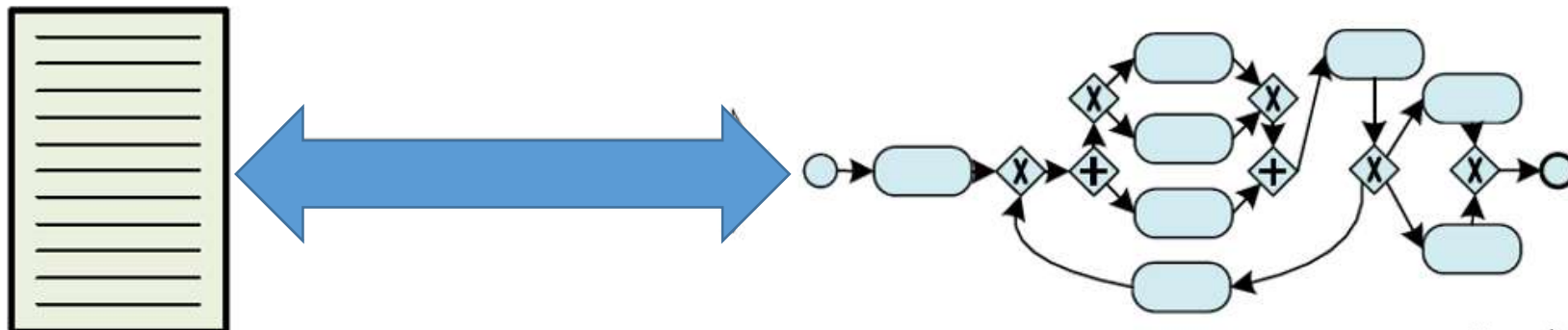


Replay

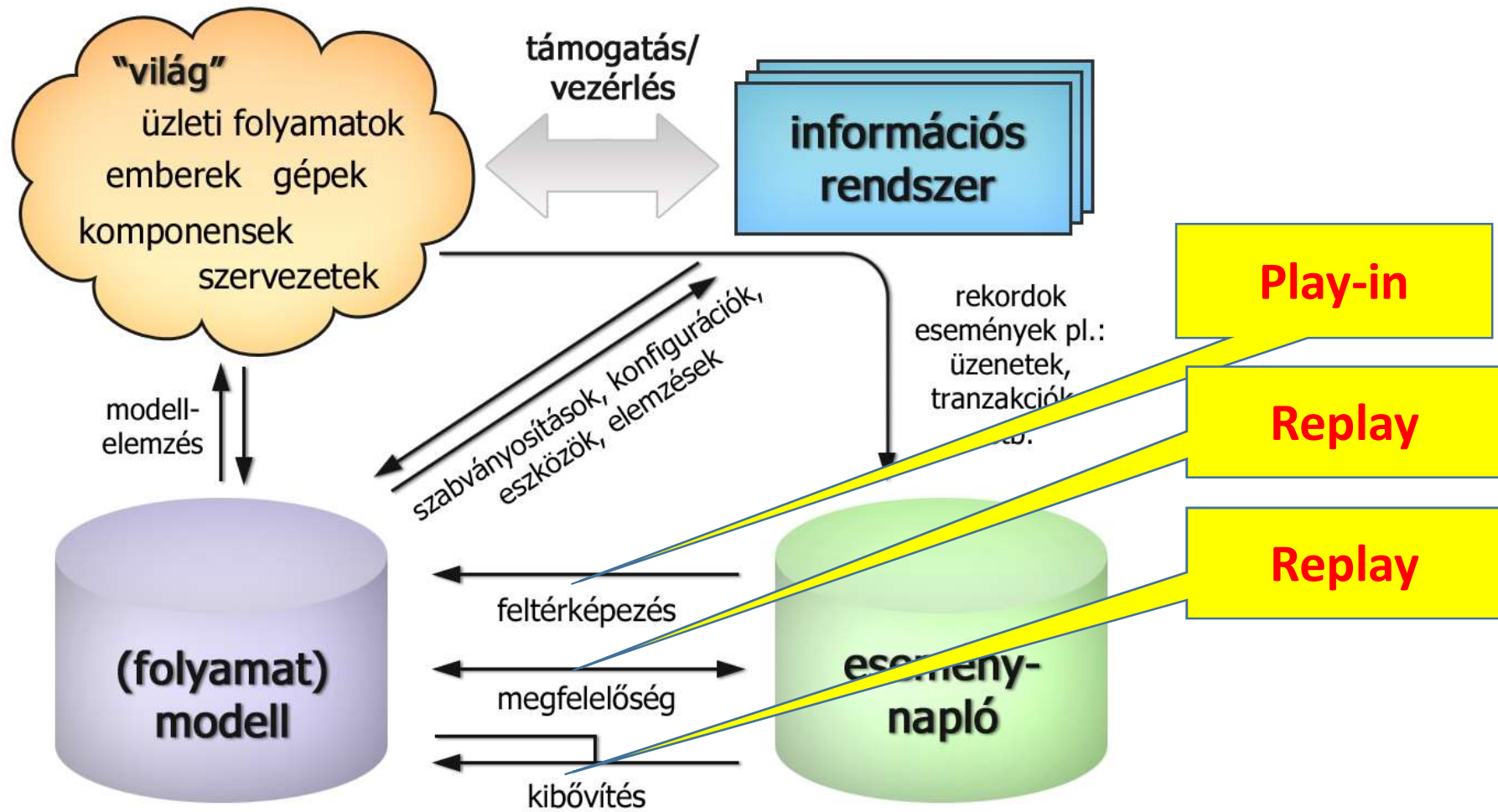
A folyamatbányászat legfontosabb formája, amihez tekintjük a modellt és a viselkedést (valós lefutásokat)  szembeállítjuk a modellt a valósággal

Használható:

- megfelelőség ellenőrzéshez (conformance checking),
- előrejelzéshez (prediction),
- torlódás elemzéshez (bottleneck analysis)
- diagnosztikához (diagnostics)
- stb.



A folyamatbányászat típusai



Esemény log (Event log)

- Feltételezzük, hogy rendelkezésre áll egy esemény napló, amelyben minden esethez tartozik egy eset azonosító, egy művelet/esemény és egy időpont.
- Az esemény log az esetek gyűjteménye.
- Egy eset az események egy szekvenciája (trace).

Az esemény adatok származhatnak:

- Egy adatbázis rendszerből (pl. betegek adatai egy kórházban),
- Egy üzleti rendszerből (pl. SAP, Oracle),
- Egy üzenet logból (pl. IBM middleware),
- Egy tranzakciós naplófájlból (pl. egy kereskedelmi rendszerből),
- Vesszővel elválasztott értékek (CSV) fájljából vagy táblázatból,
- Weboldalakról, stb.

Egy példa log

Hallgató neve	Kurzus	Vizsga időpontja	Jegy
Kovács Péter	Programozás I.	2016.01.15	4
Molnár Hajnalka	Programozás I.	2016.01.15	5
Fehér Klára	Programozás I.	2016.01.15	2
Fekete Péter	Programozás I.	2016.01.15	3
Kovács Péter	Programozás II.	2016.01.18	4
Molnár Hajnalka	Programozás II.	2016.01.18	5
Fehér Klára	Mesterséges intelligencia	2016.01.22	3
Kovács Péter	Mesterséges intelligencia	2016.01.22	5
Molnár Hajnalka	Mesterséges intelligencia	2016.01.22	4

case id

Esemény neve

Időbélyeg

Más adat

Másik példa log

order number	activity	timestamp	user	product	quantity
9901	register order	22-1-2014@09.15	Sara Jones	iPhone5S	1
9902	register order	22-1-2014@09.18	Sara Jones	iPhone5S	2
9903	register order	22-1-2014@09.27	Sara Jones	iPhone4S	1
9901	check stock	22-1-2014@09.49	Pete Scott	iPhone5S	1
9901	ship order	22-1-2014@10.11	Sue Fox	iPhone5S	1
9903	check stock	22-1-2014@10.34	Pete Scott	iPhone4S	1
9901	handle payment	22-1-2014@10.41	Carol Hope	iPhone5S	1
9902	check stock	22-1-2014@10.57	Pete Scott	iPhone5S	2
9902	cancel order	22-1-2014@11.08	Carol Hope	iPhone5S	2

case id





Esemény neve

Időbélyeg

Erőforrás

Egyéb adatok

Nem mindig világos... Számos lehetséges leképezés létezik!

- Egy email üzenet:
 - A küldő (From)  Erőforrás
 - A fogadók (To)  Egyéb adat
 - A tárgy  case id
 - Az időbélyeg (Dátum)  Időbélyeg
 - A törzs  Egyéb adat
 - etc.
- Tegyük fel, hogy egy email egy esemény.
- Mi lehet ennek a leképezése egy logban (Melyik mező a case id, melyik mező a művelet neve, stb.)?
- Problémák:
 - Nem világos mik az esetek (küldők, tárgya a leveleknek, stb.).
 - Nem világos, hogy mik a műveletek.
 - Kontextus függő, további kérdések megválaszolása szükséges!!!

Mi egy lehetséges leképezés egy esemény log-ba, mely mezők az eset azonosítók stb.?

student id	course	exam date	gender	nationality	mark
884362	Logic	15-4-2014	male	Dutch	6
884978	Logic	15-4-2014	female	Chinese	8
884362	Statistics	17-4-2014	male	Dutch	9
885087	Statistics	17-4-2014	female	German	9
882876	Statistics	17-4-2014	female	Dutch	5
..

- Tegyük fel, hogy minden sor egy eseményt azonosít.
- Egy lehetséges alternatíva pl.: a kurzus az eset, a hallgató a tevékenység és a nemzetiség az erőforrás.

Kiterjesztések

- Kiegészítő adatok egy esemény előfordulásánál:

Egy eseményhez tartozhat:

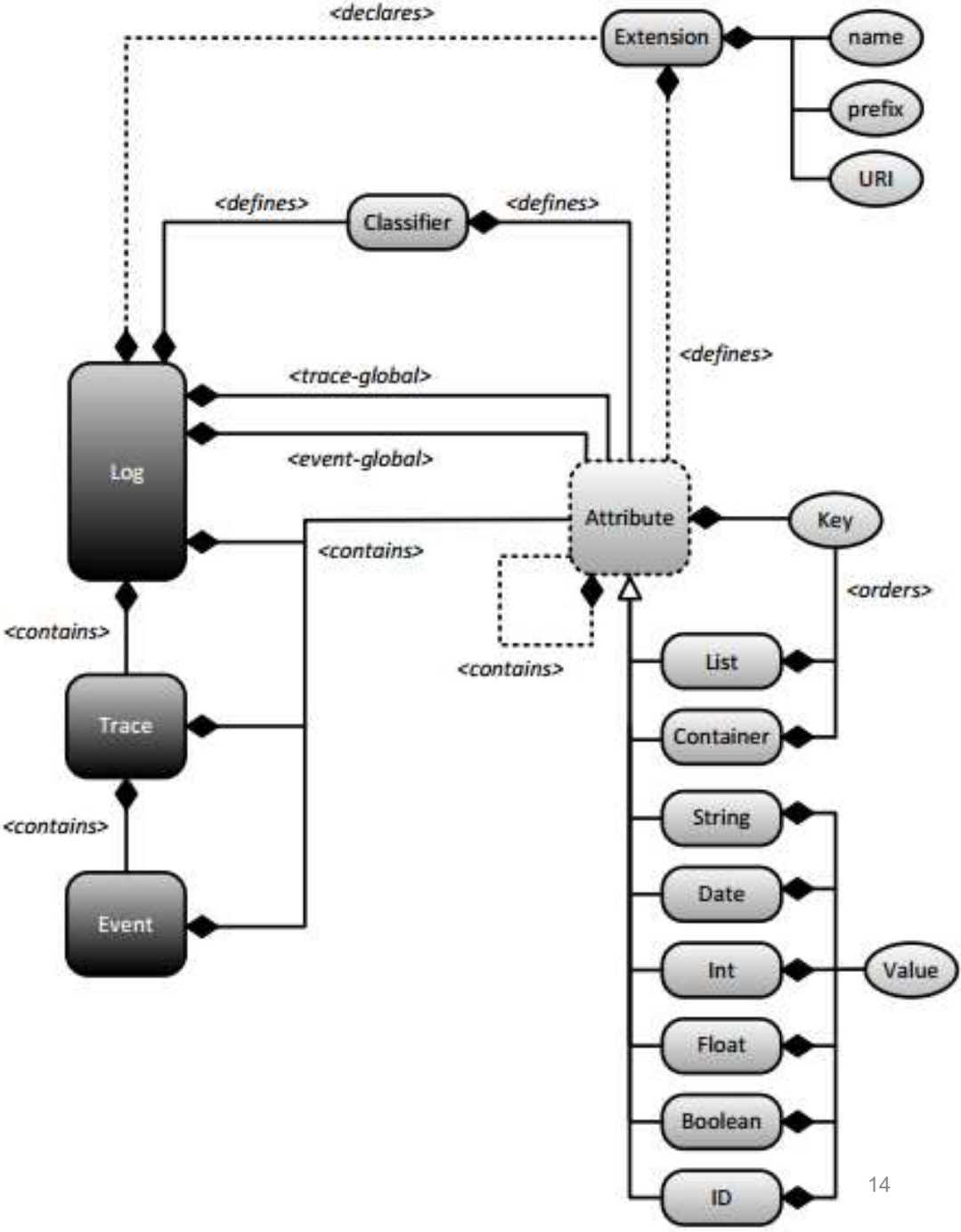
- **start** (elkezdődött),
- **complete** (teljesen végrehajtódott),
- **suspend** (fel lett függesztve a végrehajtás),
- **resume** (újra lett indítva a végrehajtás),
- **abort** (a tevékenység végrehajtása meg lett szakítva),
- stb. bejegyzés

XES (eXtensible Event Stream)

- Elfogadott a IEEE Task Force on Process Mining szervezet által.
- Ezt a formátumot támogatja a ProM és a Disco rendszer is.
- Elődök: MXML és SA-MXML.
- Könnyen lehet más formátumokra konvertálni (pl.CSV), ha a megfelelő adatok rendelkezésre állnak.
- Létezik ehhez is egy meta-modell.

XES (eXtensible Event Stream)

Meta modell



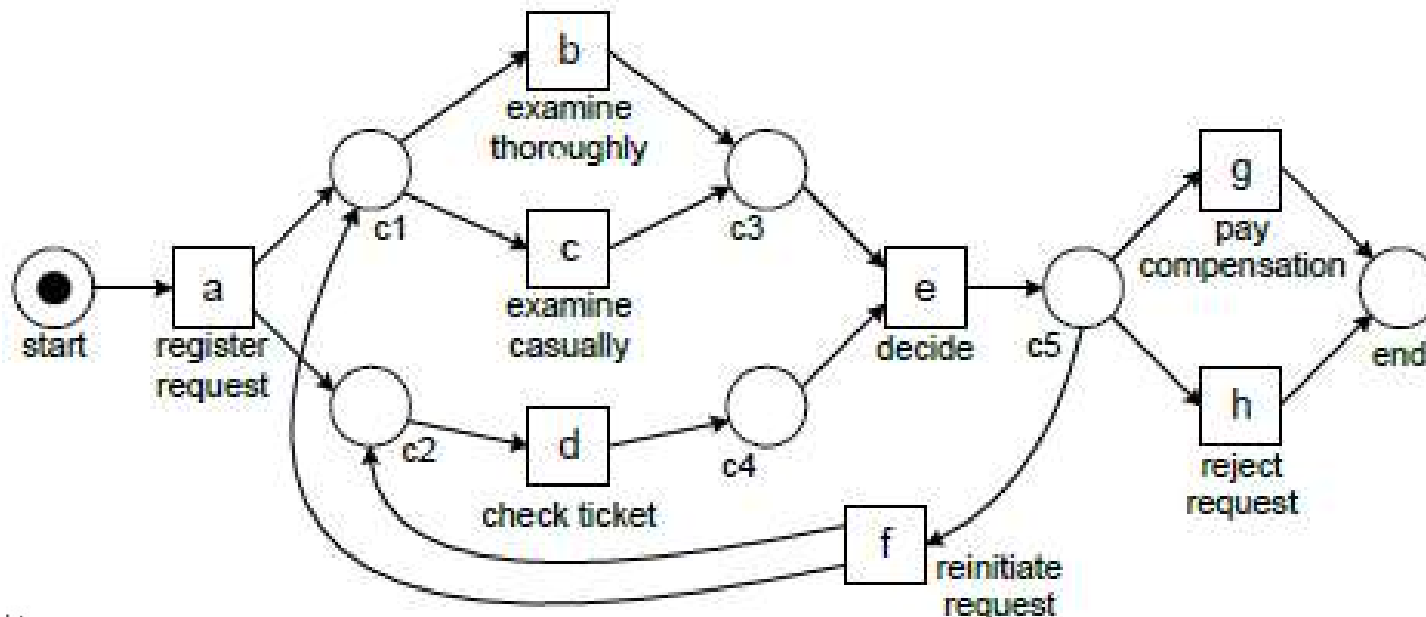
```

<?xml version="1.0" encoding="UTF-8" ?>
<log xes:version="2.0" xes:features="arbitrary-depth" xmlns="http://www.xes-standard.org
 /">
  <extension name="Concept" prefix="concept" uri="http://www.xes-standard.org/concept.
    xesext"/>
  <extension name="Time" prefix="time" uri="http://www.xes-standard.org/time.xesext"/>
  <global scope="trace">
    <string key="concept:name" value=""/>
  </global>
  <global scope="event">
    <string key="concept:name" value=""/>
    <date key="time:timestamp" value="1970-01-01T00:00:00.000+00:00"/>
    <string key="system" value=""/>
  </global>
  <classifier name="Activity" keys="concept:name"/>
  <classifier name="Another" keys="concept:name system"/>
  <float key="log attribute" value="2335.23"/>
  <trace>
    <string key="concept:name" value="Trace number one"/>
    <event>
      <string key="concept:name" value="Register client"/>
      <string key="system" value="alpha"/>
      <date key="time:timestamp" value="2009-11-25T14:12:45:000+02:00"/>
      <int key="attempt" value="23">
        <boolean key="tried hard" value="false"/>
      </int>
    </event>
    <event>
      <string key="concept:name" value="Mail rejection"/>
      <string key="system" value="beta"/>
      <date key="time:timestamp" value="2009-11-28T11:18:45:000+02:00"/>
    </event>
  </trace>
</log>

```

A „jó” reprezentáció kiválasztása! Vezérlési folyamat reprezentálása Petri háló segítségével

- Az eset kezdődik az **a** művelettel és befejeződik a **g** vagy **h** művelettel.
- A **d** művelet konkurens a **b** vagy **c** művelettel.
- Az **e** művelet várakozik mindaddig, amíg (**d** és **b**) vagy (**d** és **c**) be nem fejeződik.
- Három féle döntési lehetőségünk van: **f**, **g** vagy **h**.

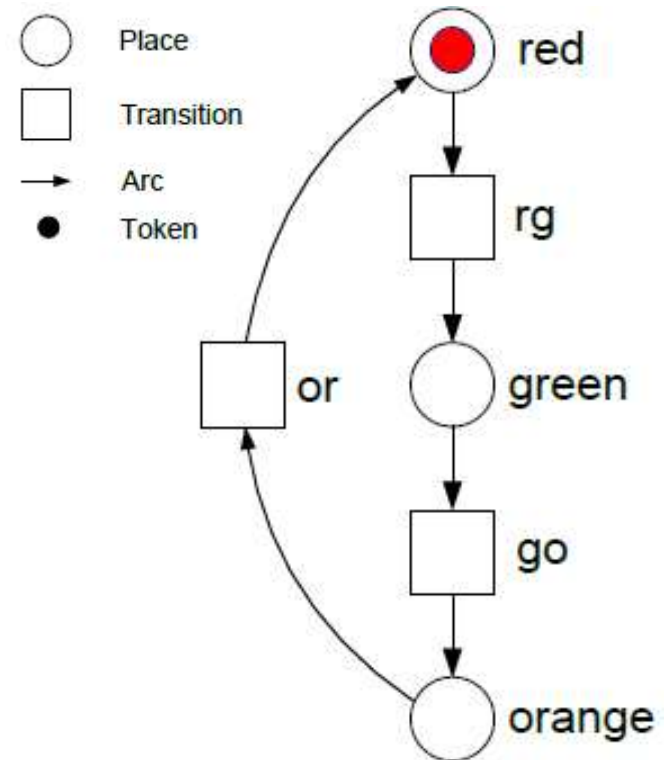
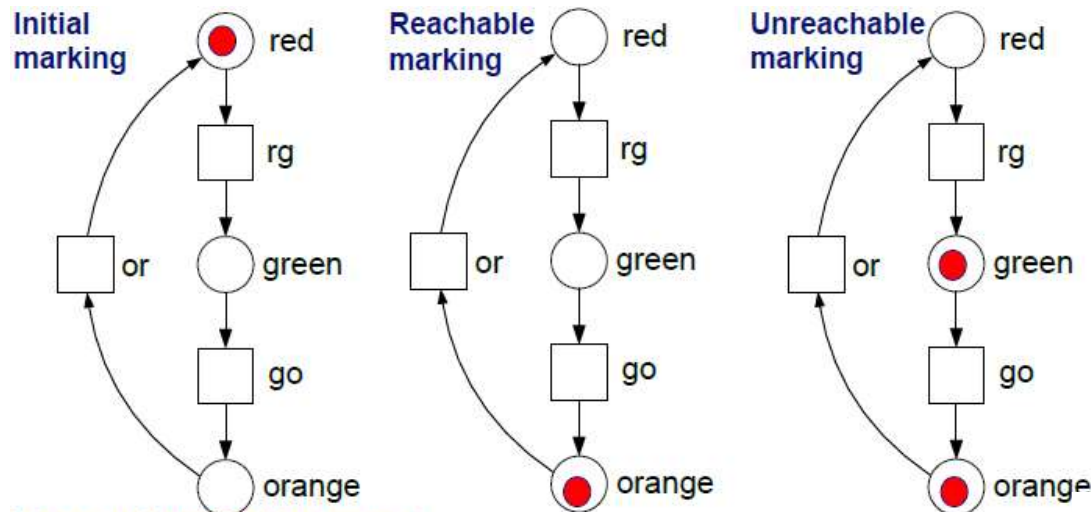


További kilátások

- Először koncentrálunk:
 - A vezérlési/irányítási folyamatokra (control-flow)
 - Adat folyamokra (data-flow)
 - Időre
 - Erőforrásokra
 - Költségekre
 - Kockázatokra
- Alternatív control-flow fogalmak:
 - BPMN (Business Process Model and Notation) diagramok
 - UML activity diagramok
 - Esemény-vezérelt folyamat láncok (Event-driven process chains)
 - Kiterjesztett Petri hálók
 - Fuzzy modellek
 - Stb.

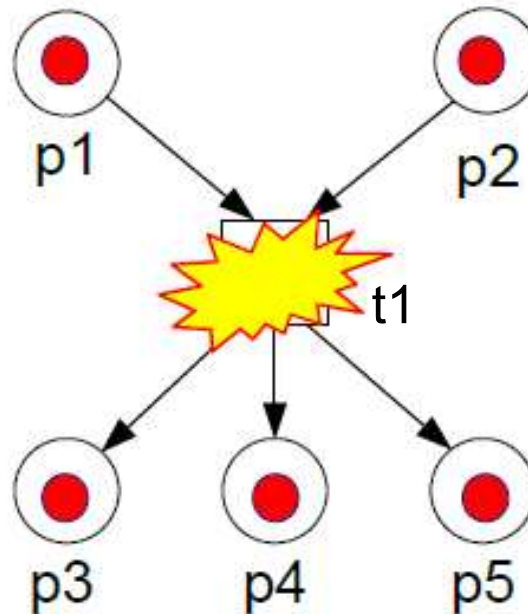
Egy egyszerű folyamat: közlekedési lámpa Petri hálója

- A hálózat állandó és tartalmaz helyeket és átmeneteket.
- A helyek tartalmazhatnak tokeneket (jelölő pontokat).
- Egy átmenet előállít és/vagy felhasznál tokeneket.

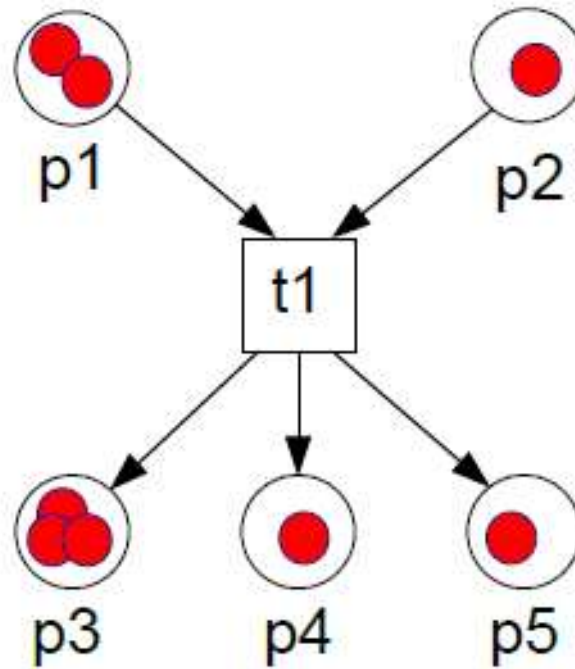


Engedélyezés és tüzelés

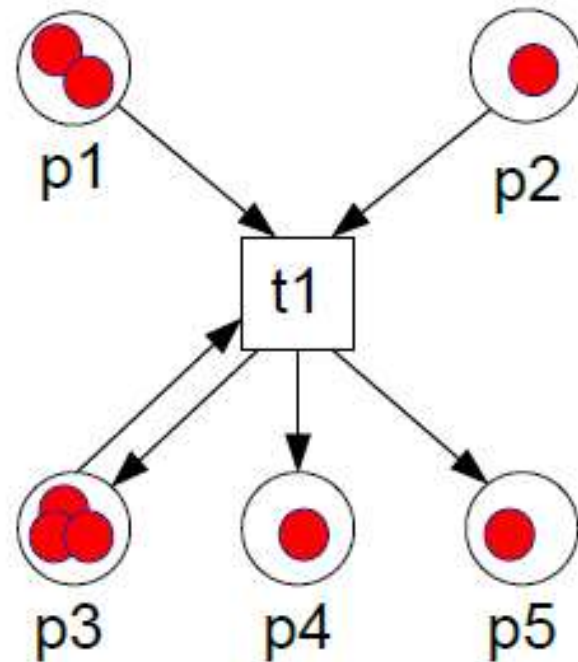
- A t1 átmenet (tranzíció) engedélyezett, ha minden input hely tartalmaz tokenet.
- Egy engedélyezett átmenet tüzel, amikor minden input helyről felhasznál egy tokenet és létrehoz minden output helyen egy tokenet.



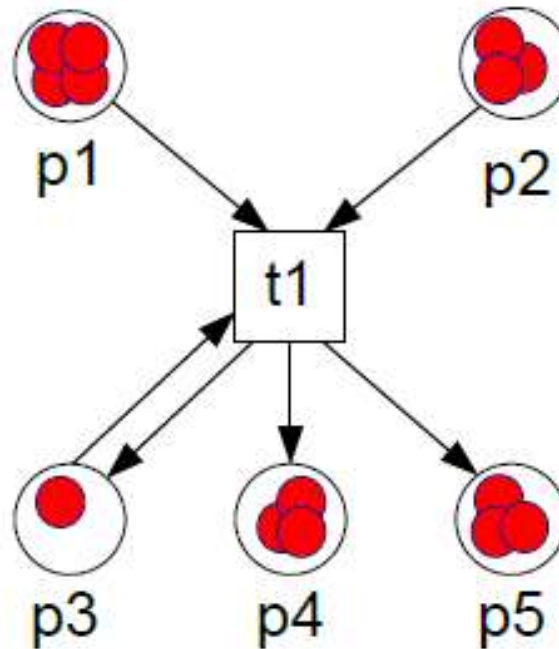
Mit eredményez az, ha egy helyen több token van?



Mi van akkor, ha egy hely egyszerre input és output hely is?

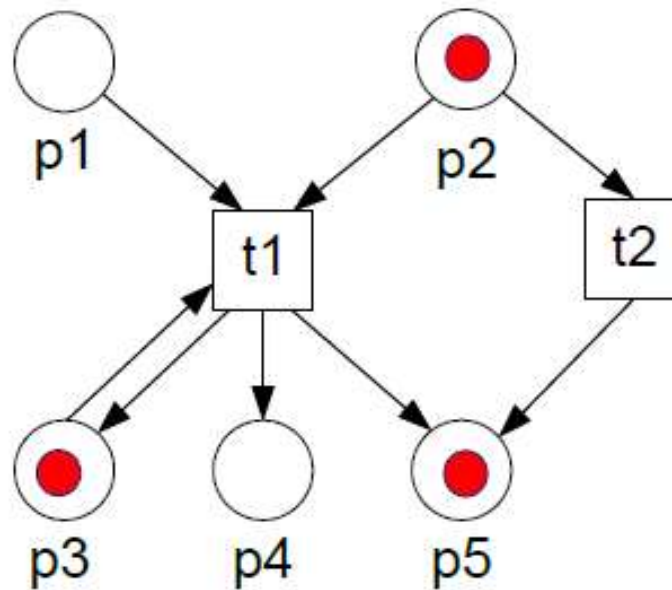


Hányszor tud a t1 tüzelni és mi lesz a végső jelölés?



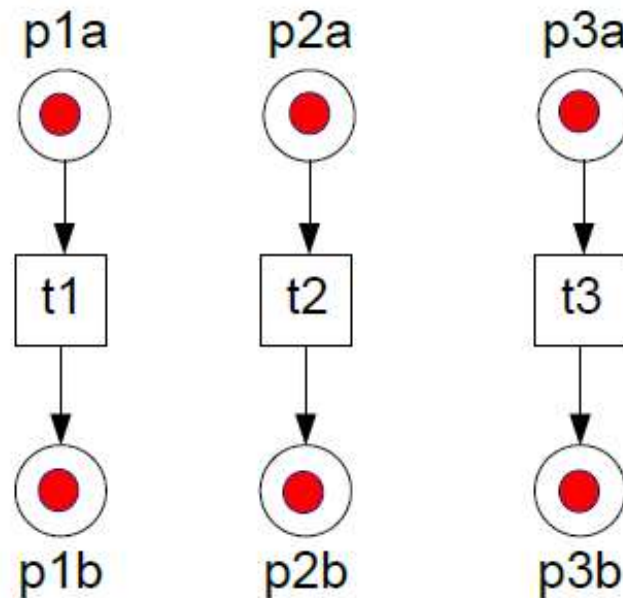
Mi történik akkor, ha több átmenet is van?

- T1 nem engedélyezett, így t2 tud tüzelni



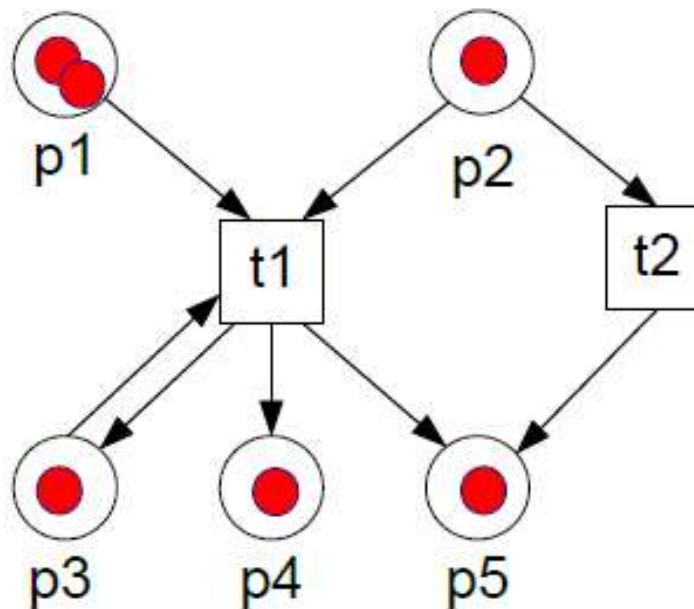
Mi történik akkor, ha több átmenet is van?

- Mindhárom átmenet engedélyezett és tud tetszőleges sorrendben vagy egyidőben tüzelni. → Párhuzamos tervezés

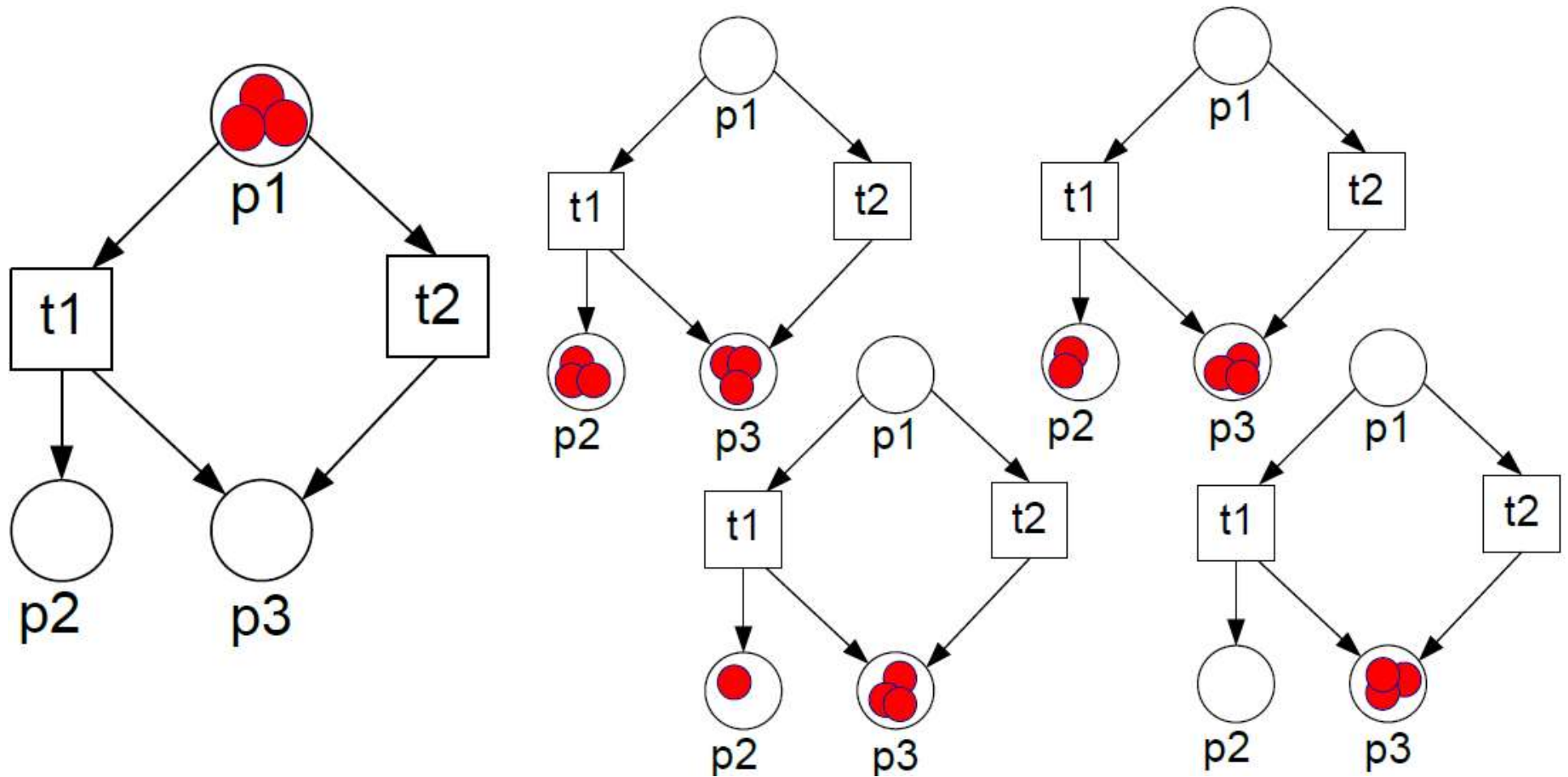


Mi történik akkor, ha több átmenet is van?

- Mindkettő engedélyezett, de csak az egyik tud tüzelni

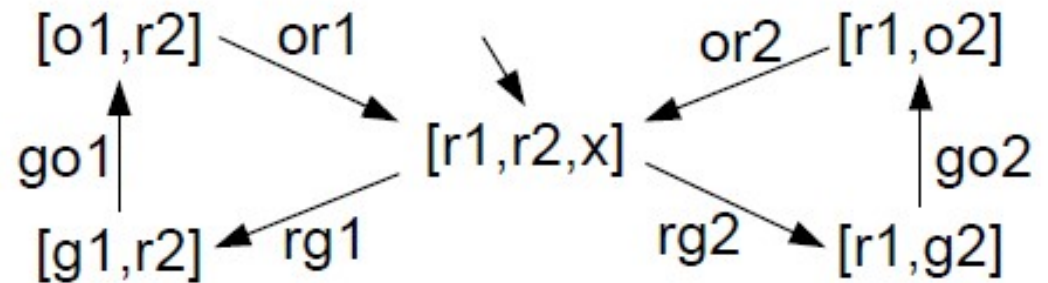
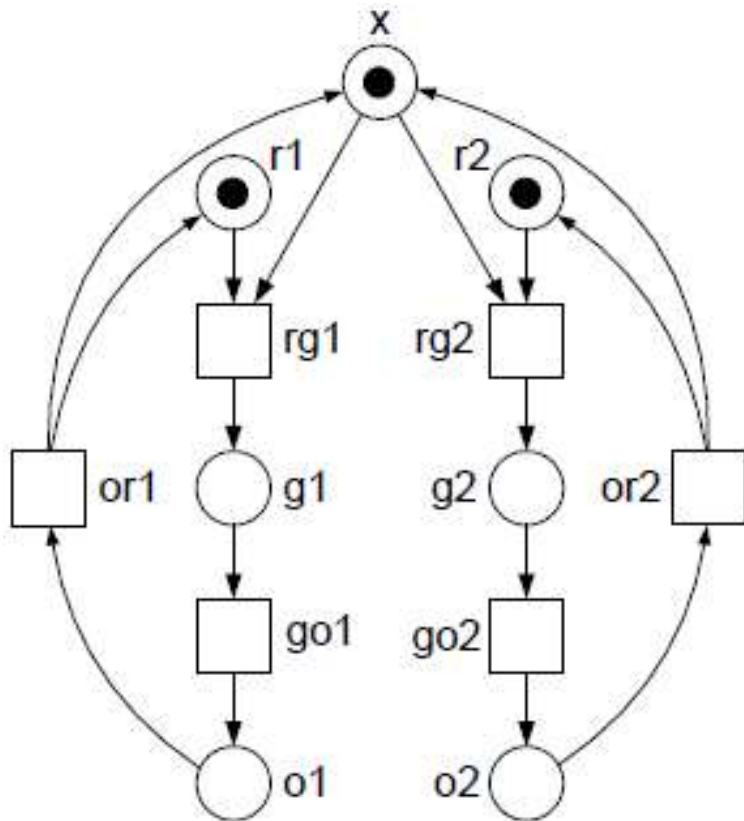


Adjuk meg az összes lehetséges vég állapot jelölést!

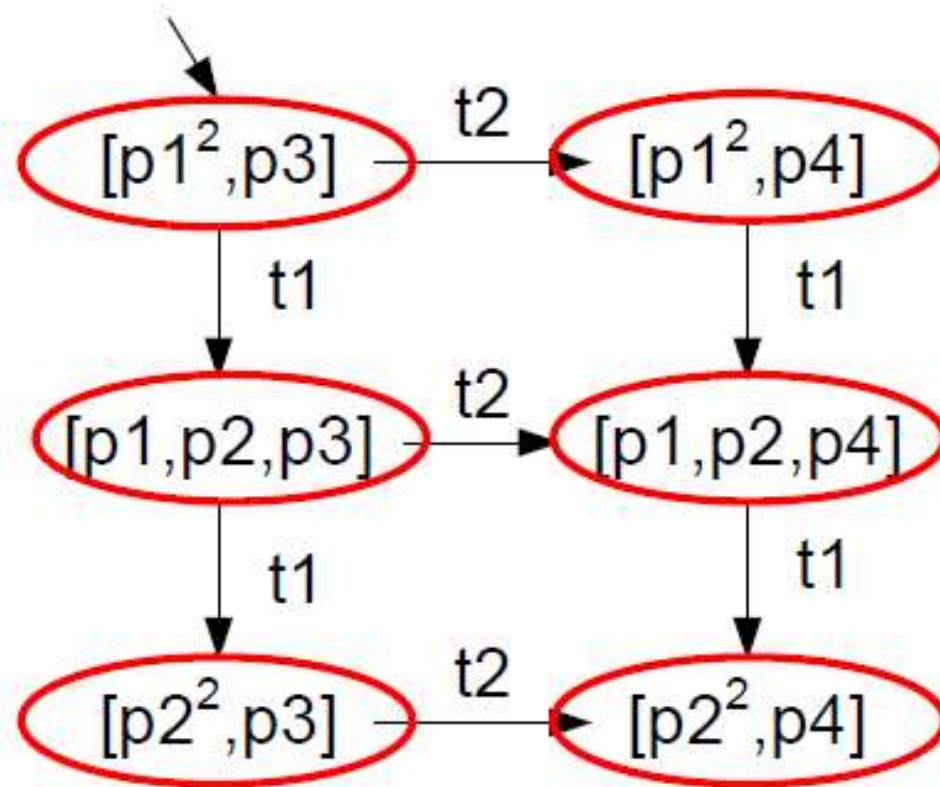
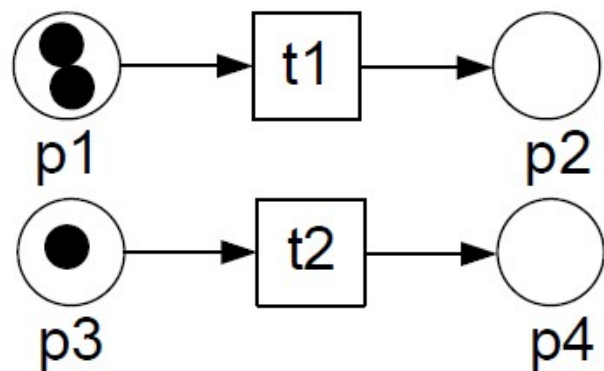


Elérhetőségi gráf

- A Petri háló elérhetőségi gráfja (reachability graph) egy átmenet rendszer (transition system) egy kezdő állapottal (kezdeti jelöléssel) és nem explicit cél állapot jelöléssel.

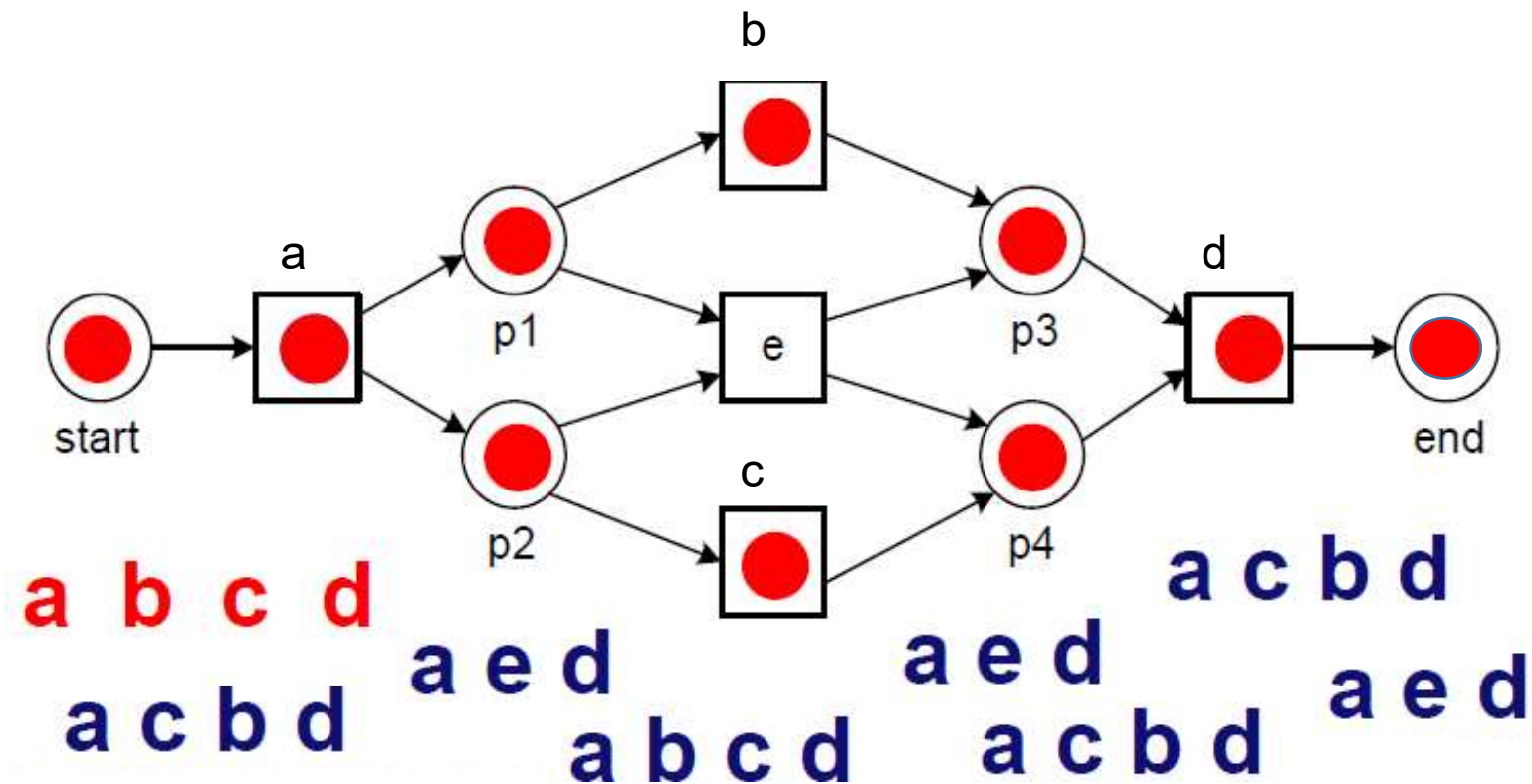


Elérhetőségi gráf



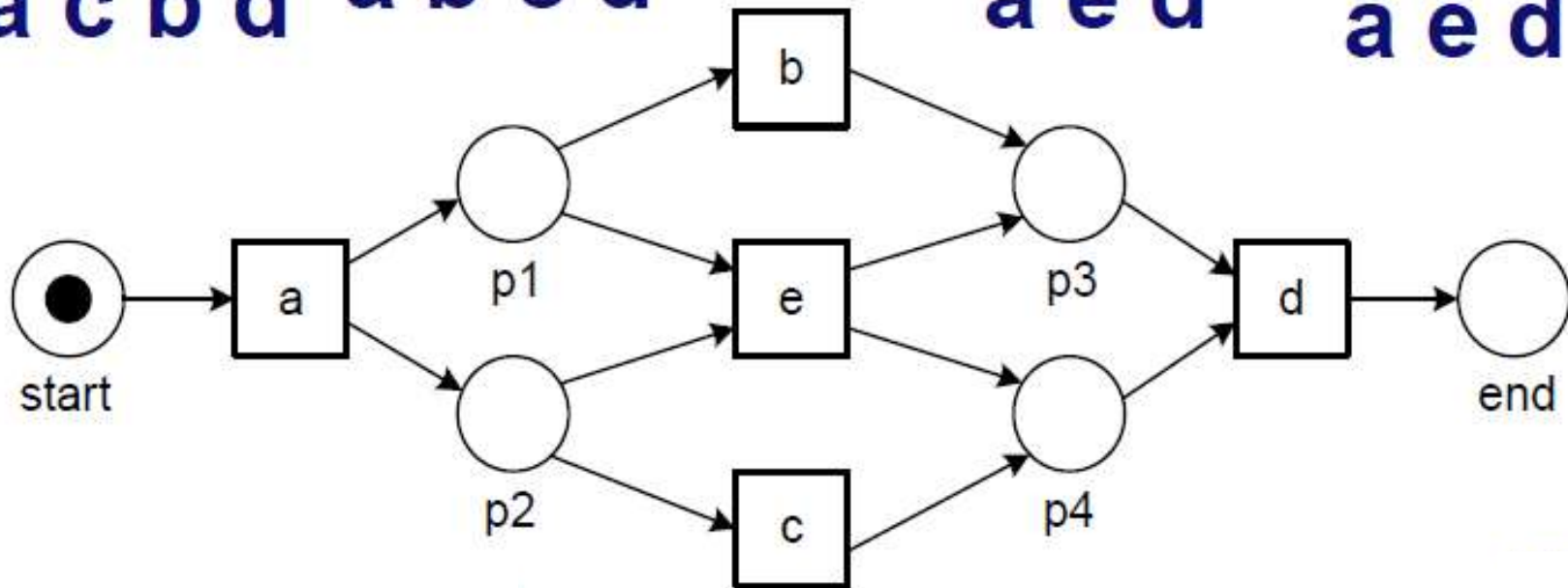
További feladatok!

A modellezés nagy kihívás!
 Play-Out (A modellek klasszikus felhasználási területe)



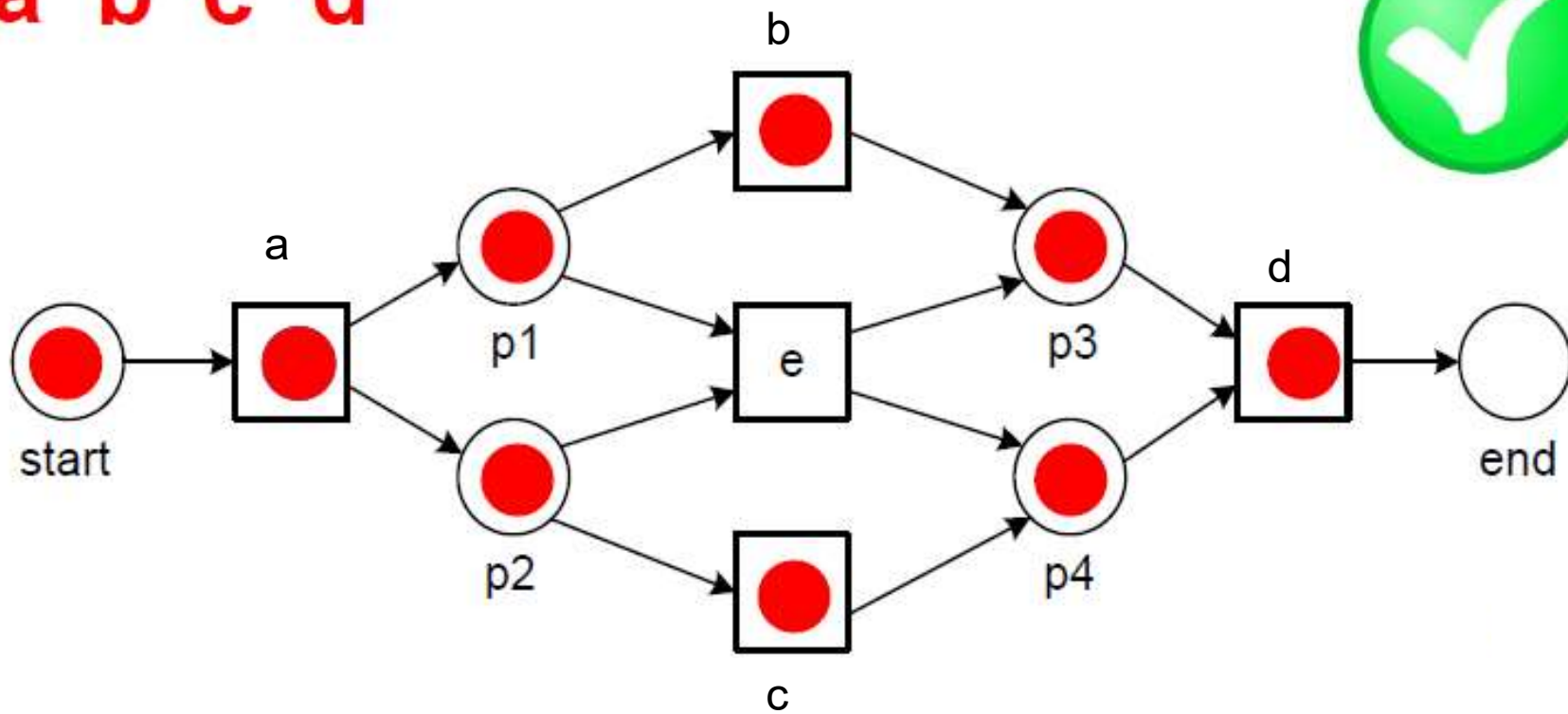
Play-In

a b c d a e d a c b d a c b d
a c b d a b c d a e d a e d



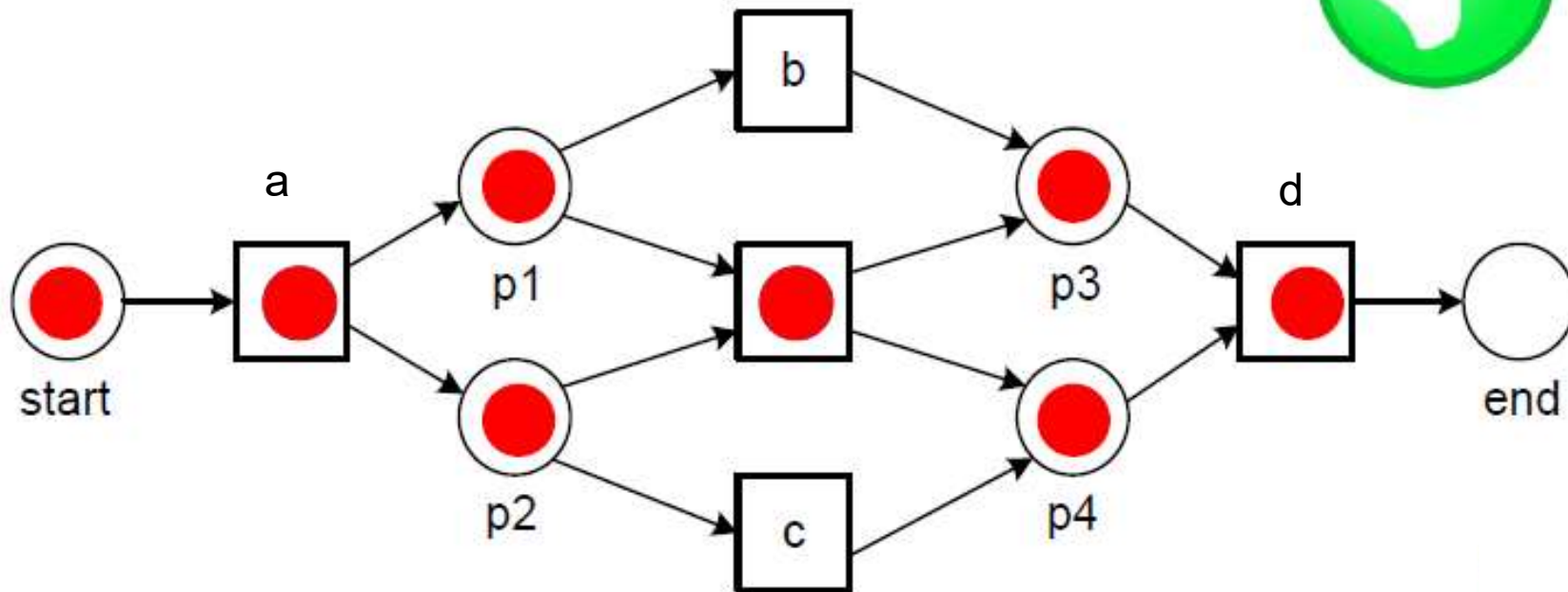
Replay

a b c d



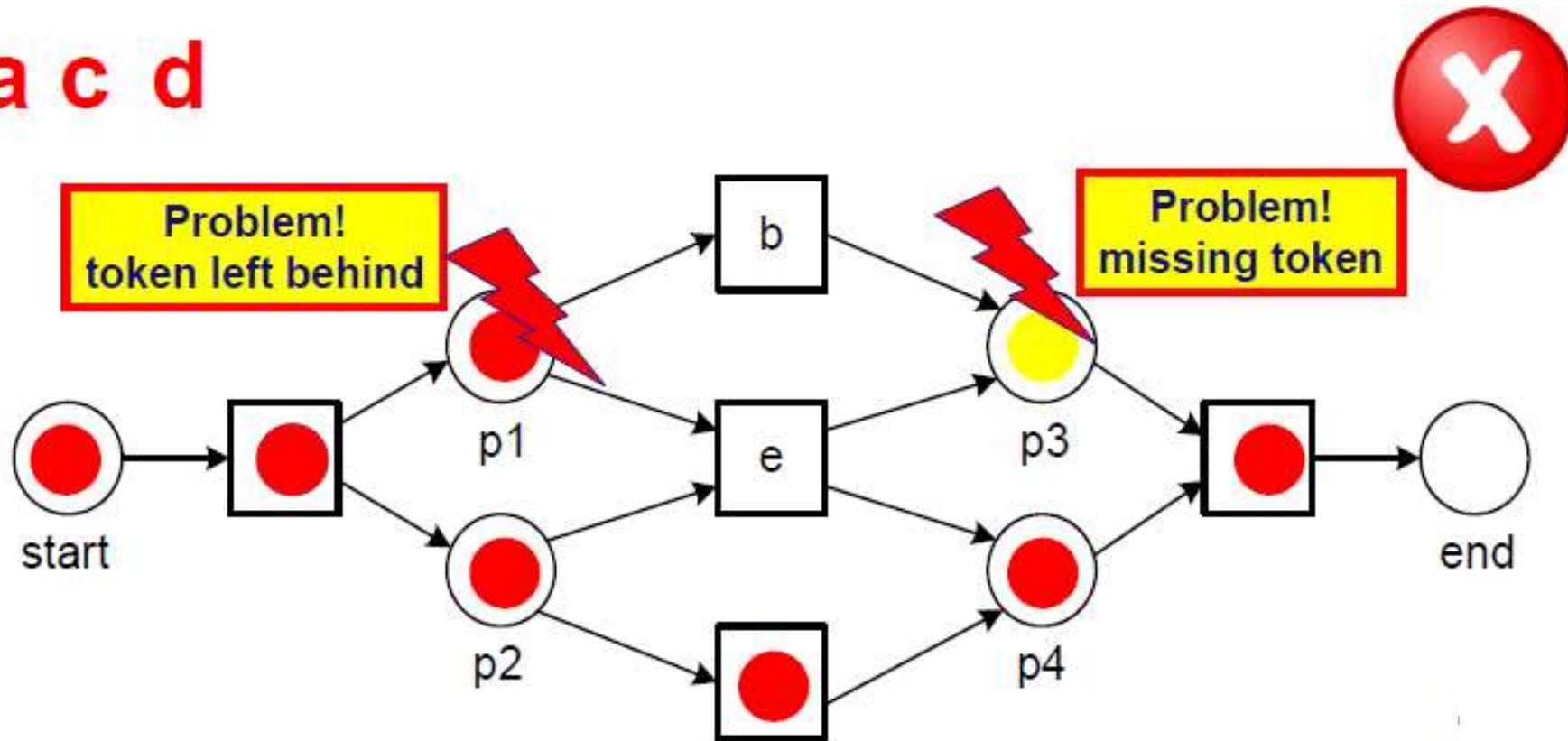
Replay

a e d



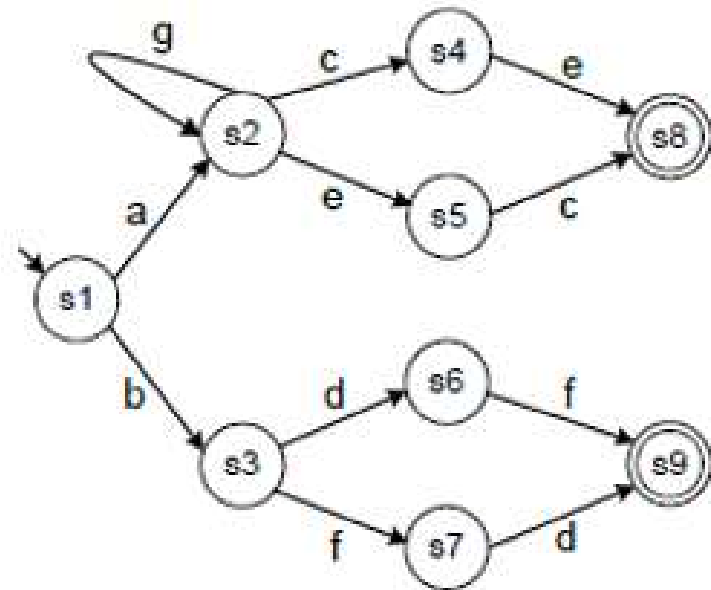
A „Replay” tevékenység azonosítani tudja a problémákat

a c d



Átmenet rendszerek (Transition systems)

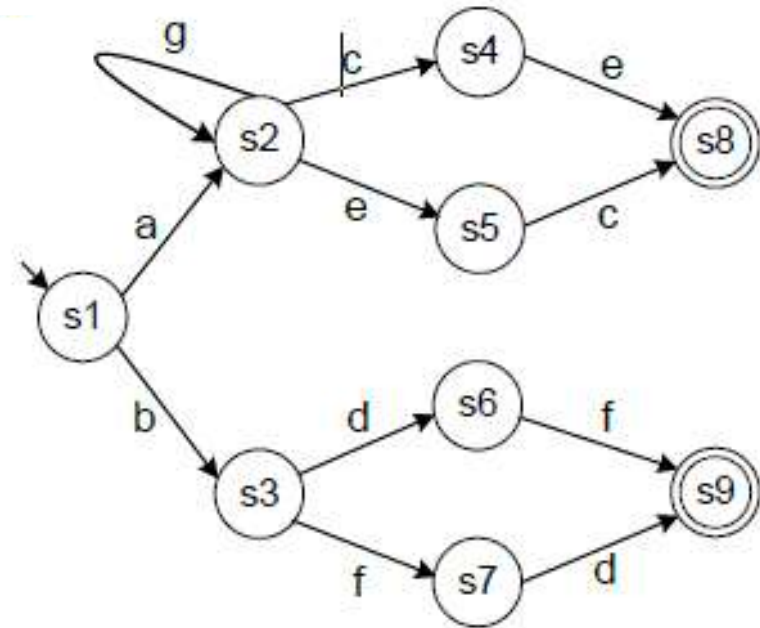
- Az elérhetőségi gráf az átmenet rendszerek egy speciális fajtája.
- A tüzelési szekvenciák megfelelnek az utaknak az átmenet rendszerben.
- Az átmenet rendszer állapotokból és átmenetekből áll.
- Lehet 1 vagy több kezdő állapot és 0 vagy több végállapot (általánosabb mint a Petri hálók).



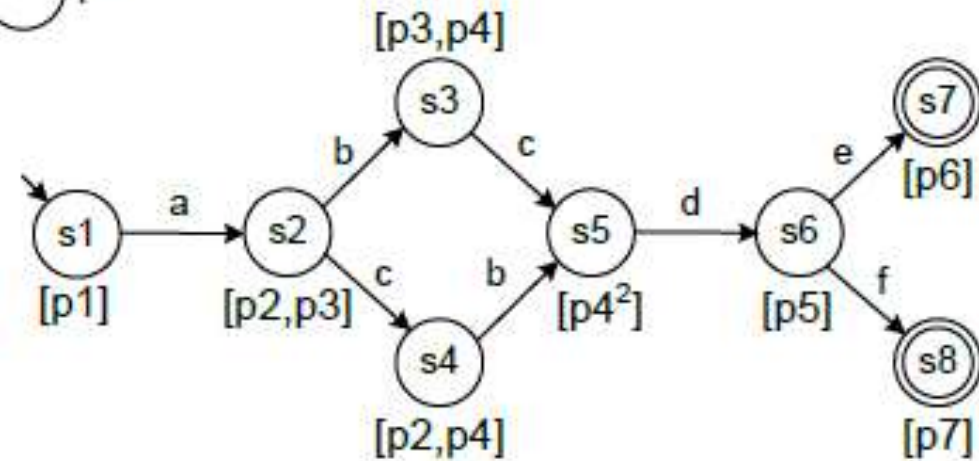
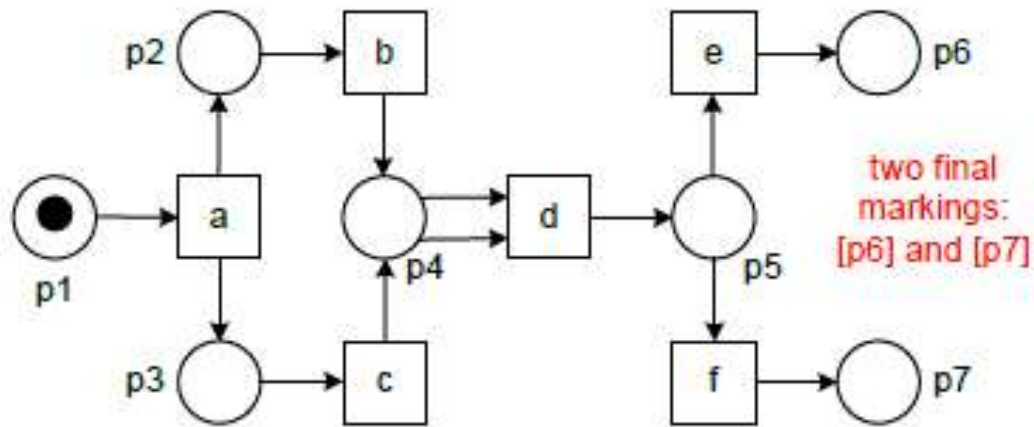
Egy kezdő állapot van: s1
Kettő végállapot: s8, s9

Teljes trace-ek

- $\langle a, g, g, c, e \rangle$ egy teljes trace.
- $\langle b, d, f \rangle$ egy teljes trace.
- $\langle a, g, g, g \rangle$ nem egy teljes (hiányos) trace.
- $\langle b, f \rangle$ nem egy teljes trace.
- Az átmenet rendszernek végtelen sok (nem) teljes trace van.

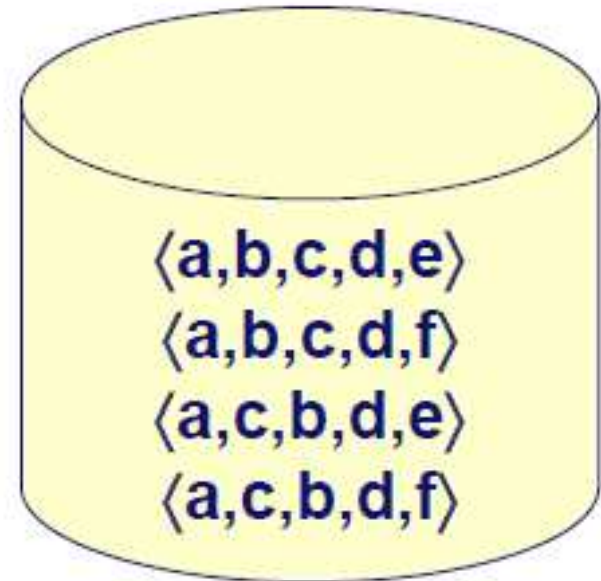
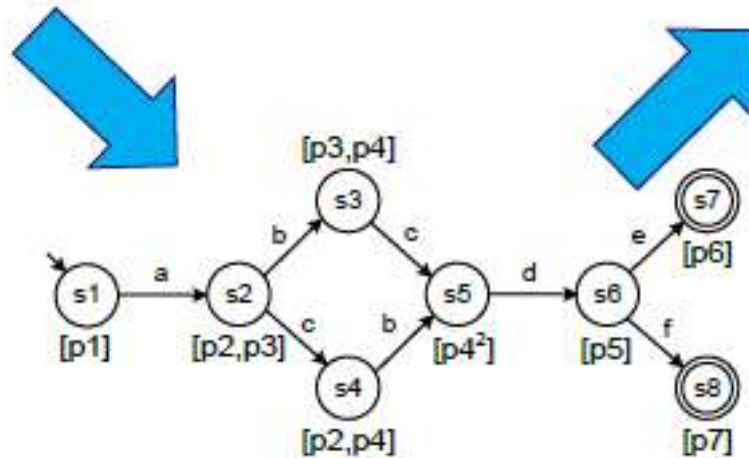
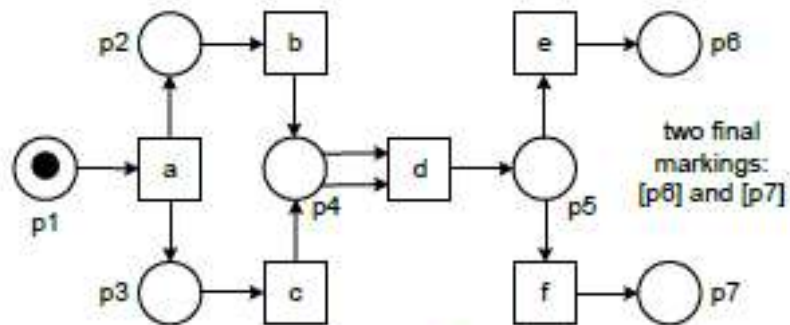


A Petri háló - Átmenet rendszer

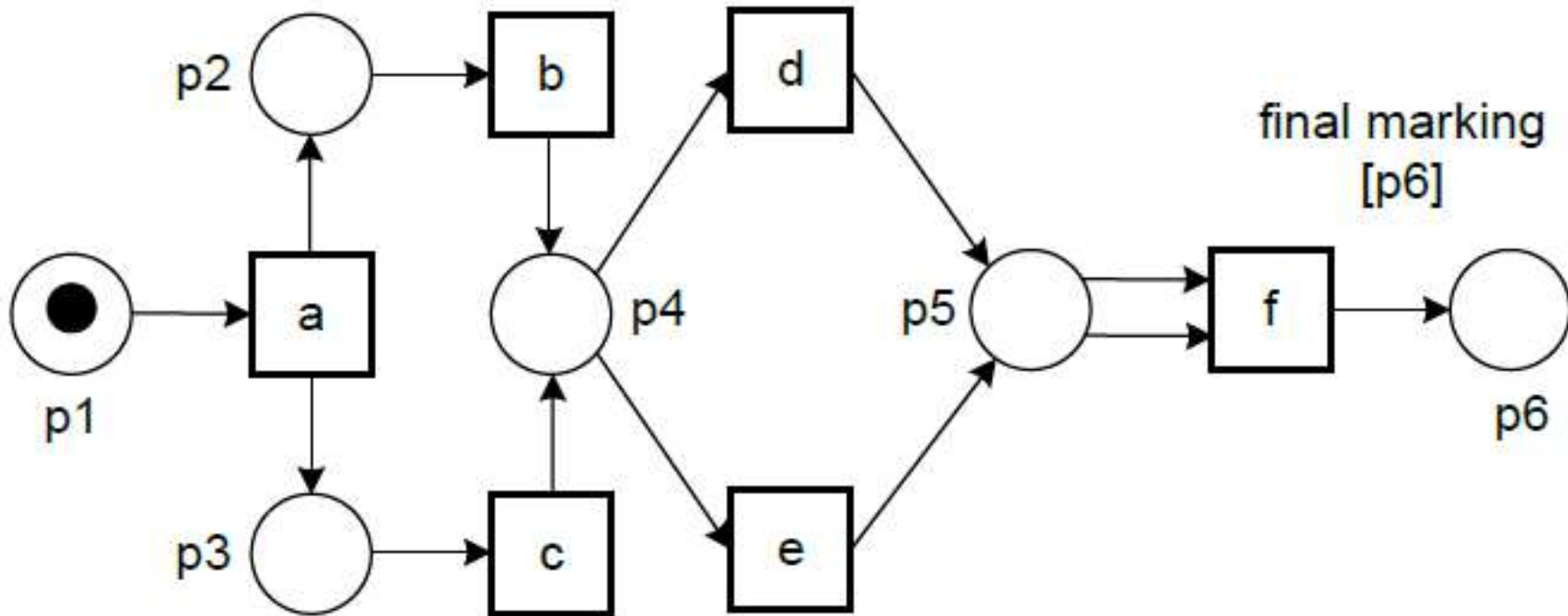


© 2004 W. van der Aalst, RWTH Aachen University

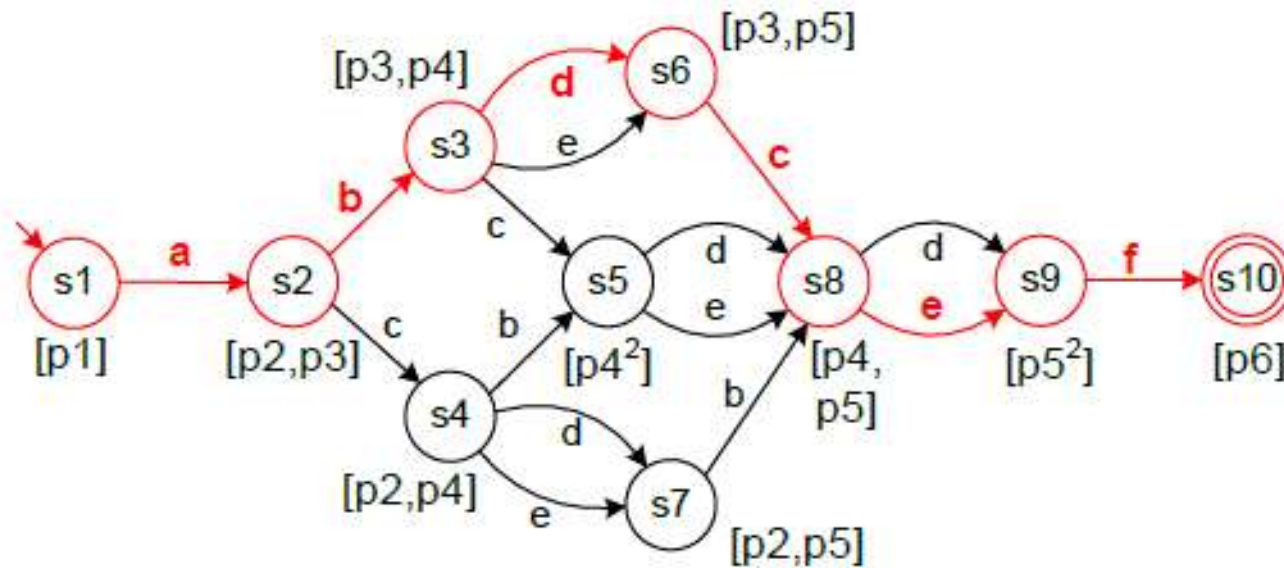
Play-Out



Play-out modell: határozza meg az átmenet rendszert és minden lehetséges teljes trace-t



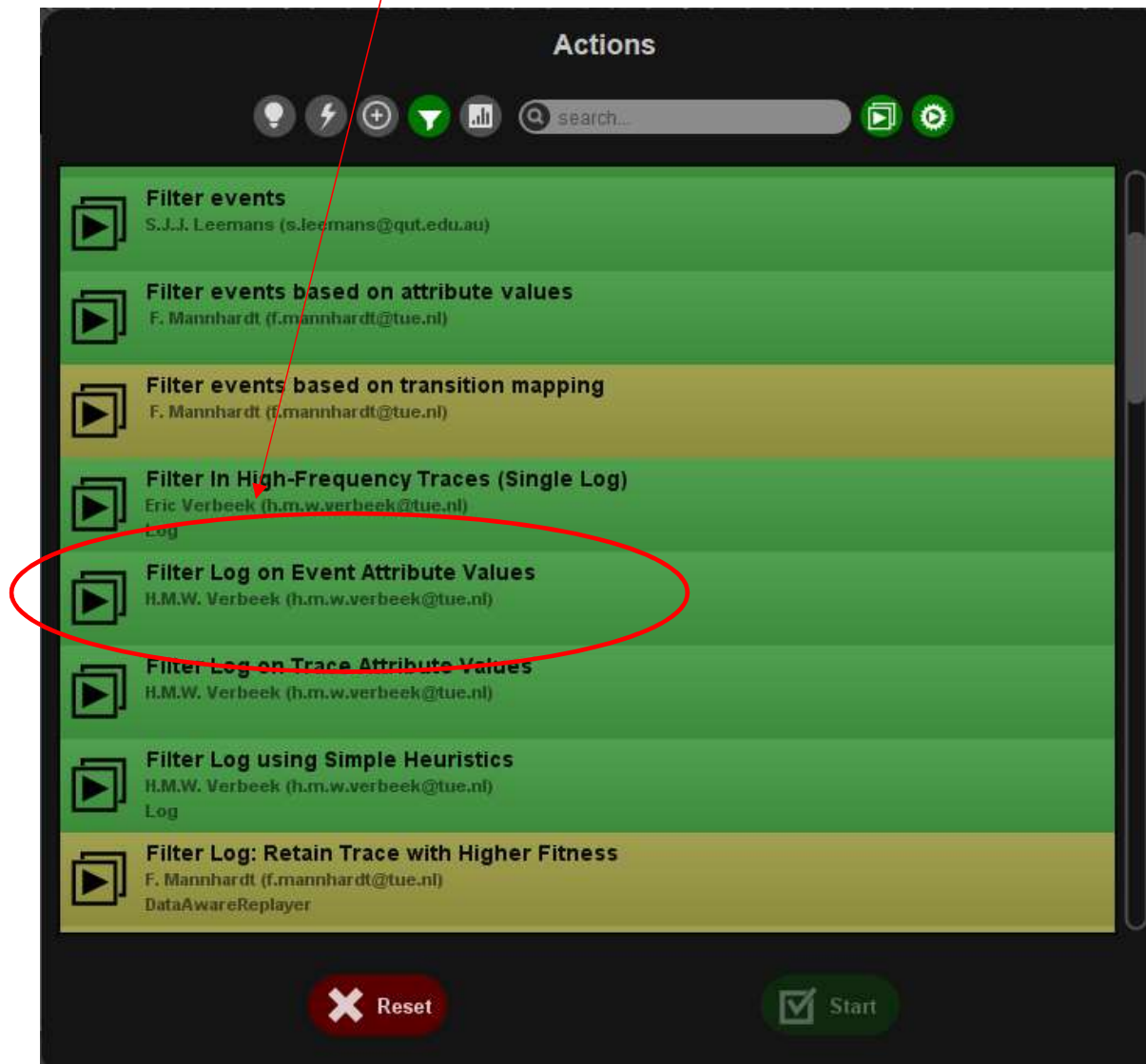
Példa egy útvonalra



- <a,b,c,d,d,f>
- <a,b,c,d,e,f>
- <a,b,c,e,d,f>
- <a,b,c,e,e,f>
- <a,c,b,d,d,f>
- <a,c,b,d,e,f>
- <a,c,b,e,d,f>
- <a,c,b,e,e,f>
- <a,b,d,c,d,f>
- <a,b,d,c,e,f>**
- <a,b,e,c,d,f>
- <a,b,e,c,e,f>
- <a,c,d,b,d,f>
- <a,c,d,b,e,f>
- <a,c,e,b,d,f>
- <a,c,e,b,e,f>

ProM használata – Log-ok tisztítása

- Miután beimportáltuk a log fájlt, különböző filtereket használhatunk, pl. Filter Log on Event Attribute Values



A kiválasztott művelet után sok féle eredményt kaphatunk, pl.:

Configure filter (values)

lifecycle:transition numberRepairs org:resource phoneType time:timestamp

concept:name defectFixed defectType

Select values

- Analyze Defect
- Archive Repair
- Inform User
- Register
- Repair (Complex)
- Repair (Simple)
- Restart Repair
- Test Repair

Remove if no value provided

Log name: DEFAULT (filtered on event attributes)

Remove trace if all events were removed

A szűrés további eredményei



Összesített adatok, az eredményeket elmenthetjük egy HTML fájlba

ProM UI Topia

ProM

DEFAULT (filtered on event attributes)

Select visualisation ...

save HTML...

Log Summary

Total number of process instances: **1104**
Total number of events: **11855**

MXML Legacy Classifier

Event classes defined by MXML Legacy Classifier

All events

Total number of classes: **12**

Class	Occurrences (absolute)	Occurrences (relative)
Test Repair+complete	1508	12,72%
Test Repair+start	1508	12,72%
Register+complete	1104	9,313%
Analyze Defect+complete	1104	9,313%
Analyze Defect+start	1104	9,313%
Inform User+complete	1102	9,296%
Archive Repair+complete	1000	8,435%
Repair (Simple)+start	785	6,622%
Repair (Simple)+complete	785	6,622%
Repair (Complex)+start	725	6,116%
Repair (Complex)+complete	724	6,107%
Restart Repair+complete	406	3,425%

Start events

Total number of classes: **1**

13:15

Átmenet rendszer – Transition system

The screenshot displays a software interface with a dark theme. At the top, there is a header bar with the word "Actions" on the left and "Activity..." on the right. Below this, the interface is divided into three main sections:

- Input:** Located on the left, it contains a tab labeled "DEFAULT" with a sub-label "XLog" and some control icons.
- Actions:** A central panel with a title "Actions" and a search bar. It lists several actions, each with a play button icon, a title, and the author's name and email:
 - Mine Pareto front with ETMd in Live mode (J.C.A.M.Buijs)
 - Mine Pareto front with ETMr (J.C.A.M.Buijs)
 - Mine Pareto front with ETMr in Live mode (J.C.A.M.Buijs)
 - Mine Transition System** (H.M.W. Verbeek) - This item is highlighted in green.
 - Multi-perspective Process Explorer - Fitness View (F. Mannhardt)
 - Multi-perspective Process Explorer - Performance View (F. Mannhardt)
 - Multi-perspective Process Explorer - Precision View (F. Mannhardt)
 - Cluster cases over time using Markov clustering (B.F.A. Hompes)
- Output:** Located on the right, it shows the results of the selected action, "Mine Transition System". The output includes:
 - Mined Transition System (TSMinerTransitionSystem)
 - Weights (DirectedGraphElementWeights)
 - Start states (StartState Set)
 - Accept states (AcceptState Set)

At the bottom of the "Actions" panel, there are two buttons: a red "Reset" button and a green "Start" button.

Beállítások

TS Miner

Configure key classifiers




Select backward keys

- MXML Legacy Classifier
- Event Name
- Resource

Select forward keys

- MXML Legacy Classifier
- Event Name
- Resource

Select key data attributes

TS Miner

Configure key classifier collections

Select collection type

- List
- Multiset
- Set
- Fixed Length Set

Select collection size limit

- No limit
- Limit: 1

Select transition system size limit

- No limit
- Limit: 200

TS Miner

Configure key classifier filter

Select 'Event Name' values

- Analyze Defect
- Archive Repair
- Inform User
- Register
- Repair (Complex)
- Repair (Simple)
- Restart Repair
- Test Repair

Select top percentage: 80

 Cancel

 Previous

 Next

TS Miner

Configure transition label filter

Select transition label values

- Analyze Defect+complete
- Analyze Defect+start
- Archive Repair+complete
- Inform User+complete
- Register+complete
- Repair (Complex)+complete
- Repair (Complex)+start
- Repair (Simple)+complete
- Repair (Simple)+start
- Restart Repair+complete
- Test Repair+complete
- Test Repair+start

Select top percentage: 80

 Cancel

 Previous

 Next

TS Miner

Configure post-mining conversions

- Remove self loops
- Improve diamond structure (may be extremely slow)
- Merge states with identical inflow
- Merge states with identical outflow
- Add artificial start and end states

 Cancel

 Previous

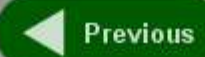
 Next

TS Miner

Check configuration

- TS Miner configuration
 - Key classifiers
 - Transition label filter
 - Post-mining conversions

 Cancel

 Previous

 Finish

PTS (mined from DEFAULT)

PIP

Zoom

