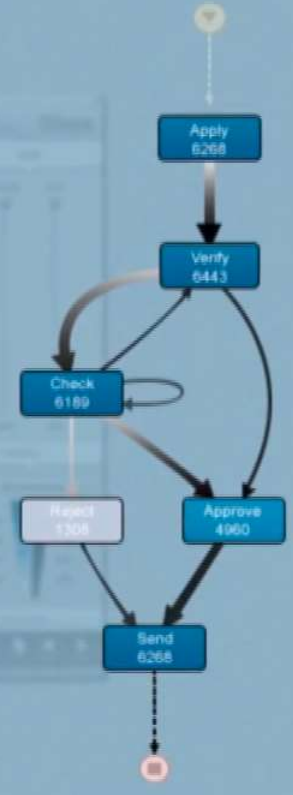
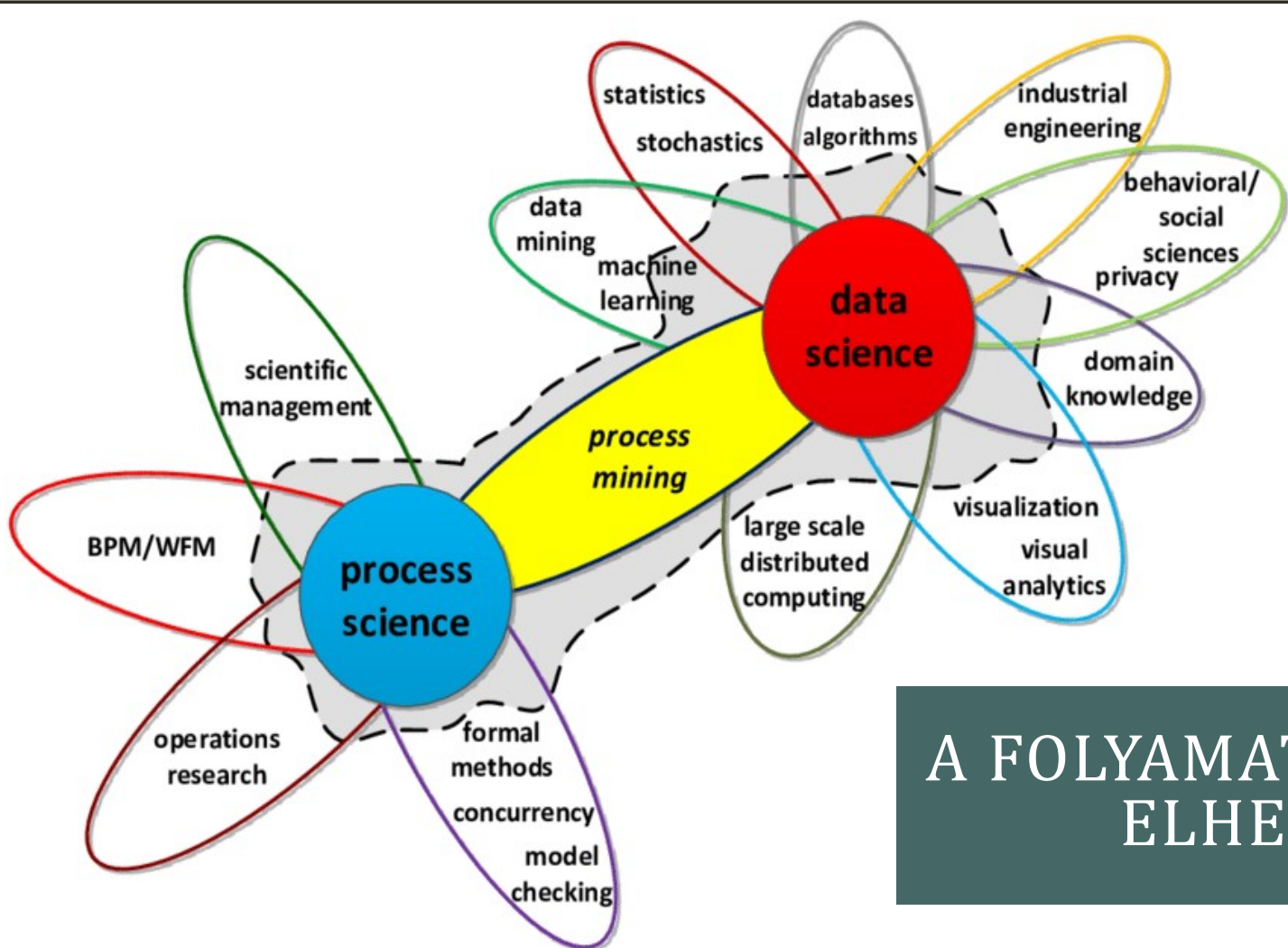


Customer ID	Activity	Start Timestamp	End Timestamp	Channel
134	Apply	2012-01-02 11:32:00.000	2012-01-02 12:02:00.000	Online
137	Apply	2012-01-02 12:19:00.000	2012-01-02 12:48:00.000	Online
110	Verify	2012-01-02 12:30:00.000	2012-01-02 12:48:00.000	Online
116	Verify	2012-01-02 13:11:00.000	2012-01-02 13:24:00.000	Online
138	Apply	2012-01-02 13:19:00.000	2012-01-02 13:28:00.000	Online
119	Verify	2012-01-02 13:26:00.000	2012-01-02 14:46:00.000	Online
118	Verify	2012-01-02 13:28:00.000	2012-01-02 14:20:00.000	Online
117	Verify	2012-01-02 13:30:00.000	2012-01-02 14:28:00.000	Online
139	Apply	2012-01-02 13:33:00.000	2012-01-02 13:51:00.000	Online
140	Apply	2012-01-02 13:33:00.000	2012-01-02 13:56:00.000	Online
141	Apply	2012-01-02 13:35:00.000	2012-01-02 13:54:00.000	Online
142	Apply	2012-01-02 13:39:00.000	2012-01-02 13:58:00.000	Online
124	Verify	2012-01-02 14:18:00.000	2012-01-02 14:21:00.000	Online
144	Apply	2012-01-02 14:22:00.000	2012-01-02 14:33:00.000	Online
122	Verify	2012-01-02 14:30:00.000	2012-01-02 15:44:00.000	Online
145	Apply	2012-01-02 14:33:00.000	2012-01-02 15:07:00.000	Online
147	Apply	2012-01-02 14:36:00.000	2012-01-02 15:03:00.000	Online
120	Verify	2012-01-02 14:39:00.000	2012-01-02 15:07:00.000	Online
148	Apply	2012-01-02 14:47:00.000	2012-01-02 15:01:00.000	Online
129	Verify	2012-01-02 14:58:00.000	2012-01-02 16:03:00.000	Online
143	Apply	2012-01-02 14:47:00.000	2012-01-02 16:13:00.000	Local Branch
127	Verify	2012-01-02 15:33:00.000	2012-01-02 16:47:00.000	Online
153	Apply	2012-01-02 16:10:00.000	2012-01-02 16:44:00.000	Local Branch
135	Verify	2012-01-02 16:37:00.000	2012-01-02 17:50:00.000	Online
154	Apply	2012-01-02 17:00:00.000	2012-01-02 17:33:00.000	Local Branch
132	Verify	2012-01-02 17:08:00.000	2012-01-02 17:58:00.000	Online
168	Apply	2012-01-02 17:28:00.000	2012-01-02 17:59:00.000	Local Branch
157	Apply	2012-01-02 17:28:00.000	2012-01-02 17:44:00.000	Local Branch
155	Apply	2012-01-02 17:44:00.000	2012-01-02 18:01:00.000	Local Branch
126	Apply	2012-01-02 18:08:00.000	2012-01-02 18:21:00.000	Local Branch
161	Apply	2012-01-02 18:07:00.000	2012-01-02 18:49:00.000	Local Branch
162	Apply	2012-01-02 18:32:00.000	2012-01-02 18:52:00.000	Local Branch
134	Verify	2012-01-02 18:44:00.000	2012-01-02 18:52:00.000	Online
163	Apply	2012-01-02 18:53:00.000	2012-01-02 19:10:00.000	Local Branch
165	Apply	2012-01-02 18:53:00.000	2012-01-02 19:33:00.000	Local Branch
166	Apply	2012-01-02 18:54:00.000	2012-01-02 20:00:00.000	Local Branch
137	Verify	2012-01-02 18:52:00.000	2012-01-02 20:04:00.000	Online
169	Apply	2012-01-02 19:01:00.000	2012-01-02 20:27:00.000	Local Branch
136	Verify	2012-01-02 20:08:00.000	2012-01-02 21:18:00.000	Online
138	Verify	2012-01-02 20:10:00.000	2012-01-02 20:38:00.000	Online
142	Verify	2012-01-02 20:10:00.000	2012-01-02 20:39:00.000	Online
141	Verify	2012-01-02 20:16:00.000	2012-01-02 21:03:00.000	Online
140	Verify	2012-01-02 20:22:00.000	2012-01-02 20:47:00.000	Online
171	Apply	2012-01-02 21:00:00.000	2012-01-02 21:19:00.000	Local Branch
147	Verify	2012-01-02 21:28:00.000	2012-01-02 22:09:00.000	Online
144	Verify	2012-01-02 21:30:00.000	2012-01-02 22:04:00.000	Online
151	Verify	2012-01-02 21:47:00.000	2012-01-02 22:07:00.000	Local Branch



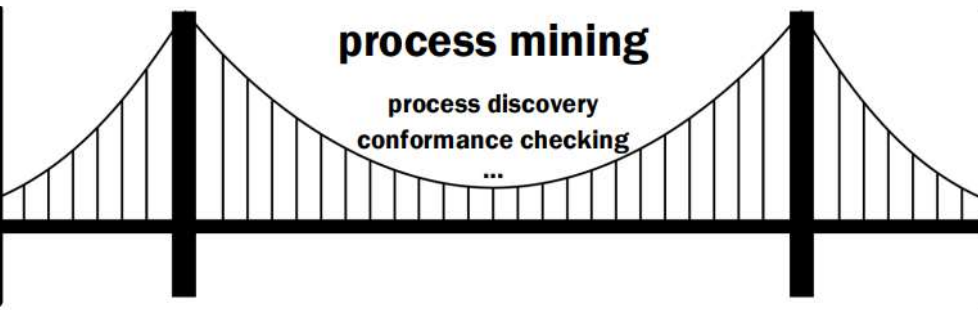
FOLYAMATBÁNYÁSZAT 3

A folyamatbányászat különböző típusai,
munkafolyamatok elemzése, Alfa algoritmus



A FOLYAMATBÁNYÁSZAT ELHELYEZKEDÉSE

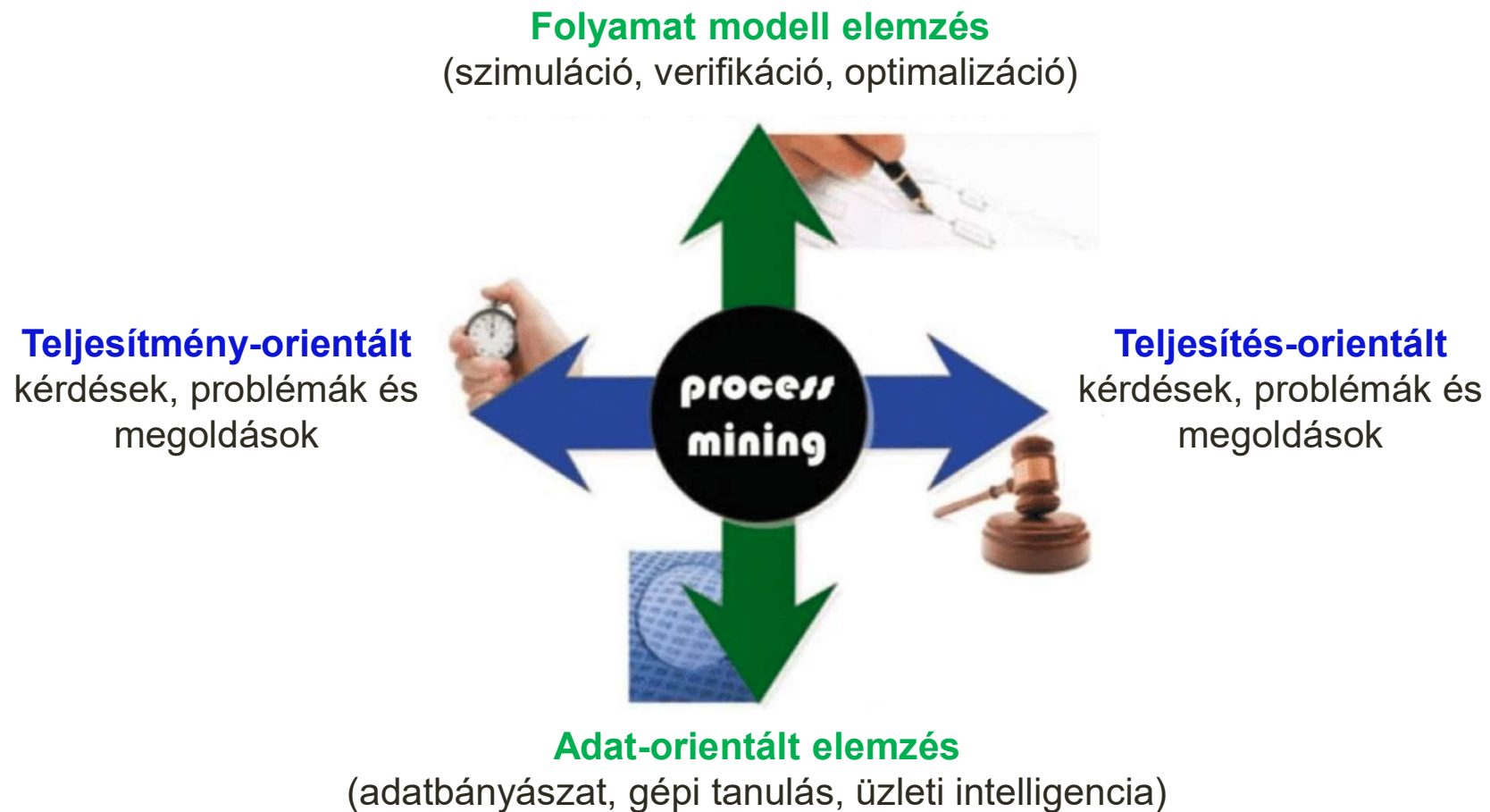
- process science**
- workflow management
 - simulation
 - business process reengineering
 - scientific management
 - business process management
 - verification
 - operations research
 - concurrency theory



- data science**
- NoSQL
 - machine learning
 - predictive analytics
 - visual analytics
 - data mining
 - clustering
 - MapReduce
 - statistics
 - ethics

Diagram by W. van der Aalst

A FOLYAMATBÁNYÁSZAT HELYMEGHATÁROZÁSA



KIINDULÁSI PONT A FOLYAMATBÁNYÁSZATHOZ: ESEMÉNY ADATOK

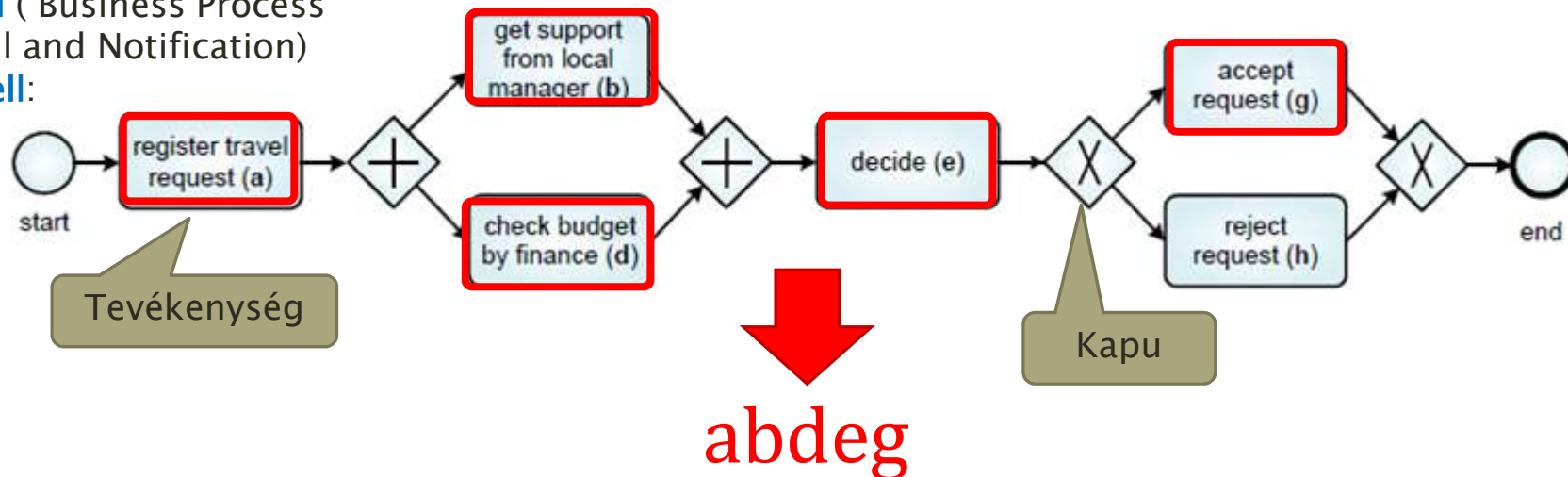
Eset azonosító Case ID	Esemény Activity	Időbélyeg Timestamp	Más adat Attribute
Hallgató neve	Kurzus	Vizsga időpontja	Jegy
Kovács Péter	Programozás I.	2016.01.15	4
Molnár Hajnalka	Programozás I.	2016.01.15	5
Fehér Klára	Programozás I.	2016.01.15	2
Fekete Péter	Programozás I.	2016.01.15	3
Kovács Péter	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

1 eset

...

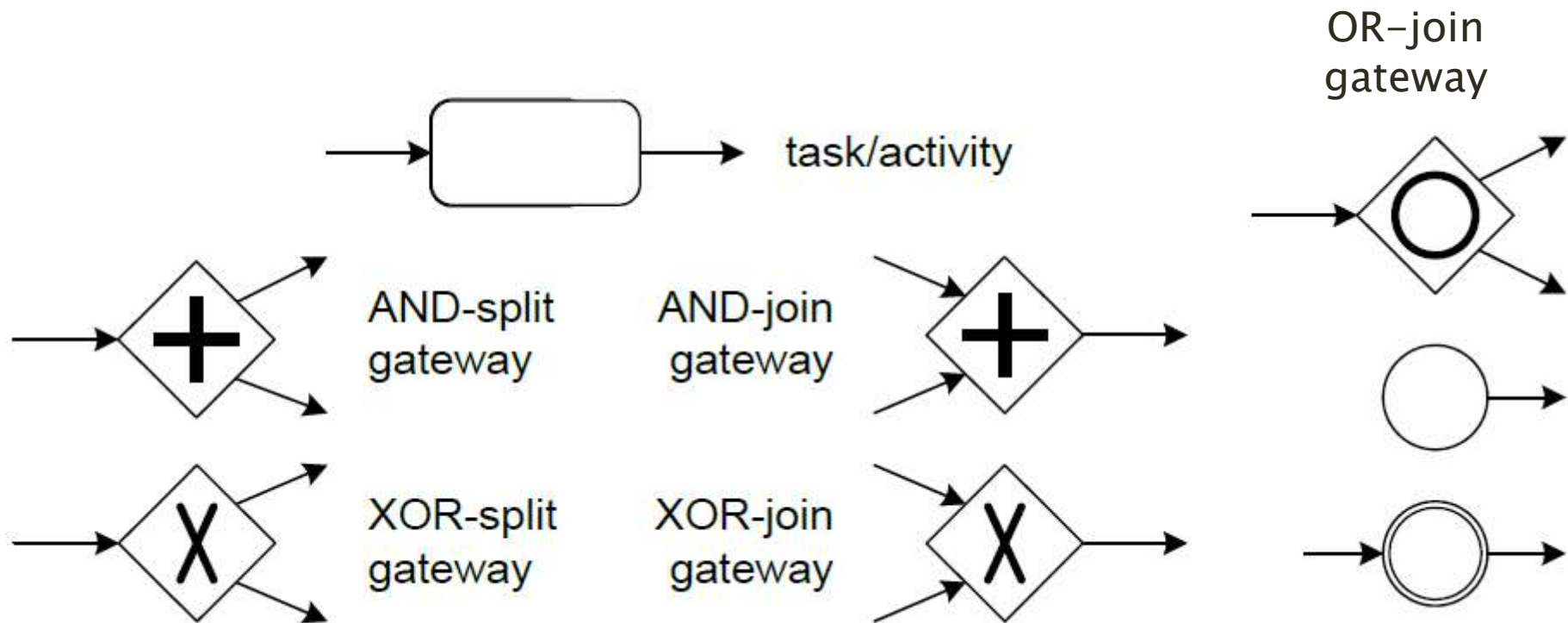
PLAY-OUT: MODEL → LEFUTÁSOK (LOG) EGY LEHETSÉGES FORGATÓKÖNYV

BPMN (Business Process Model and Notification) modell:

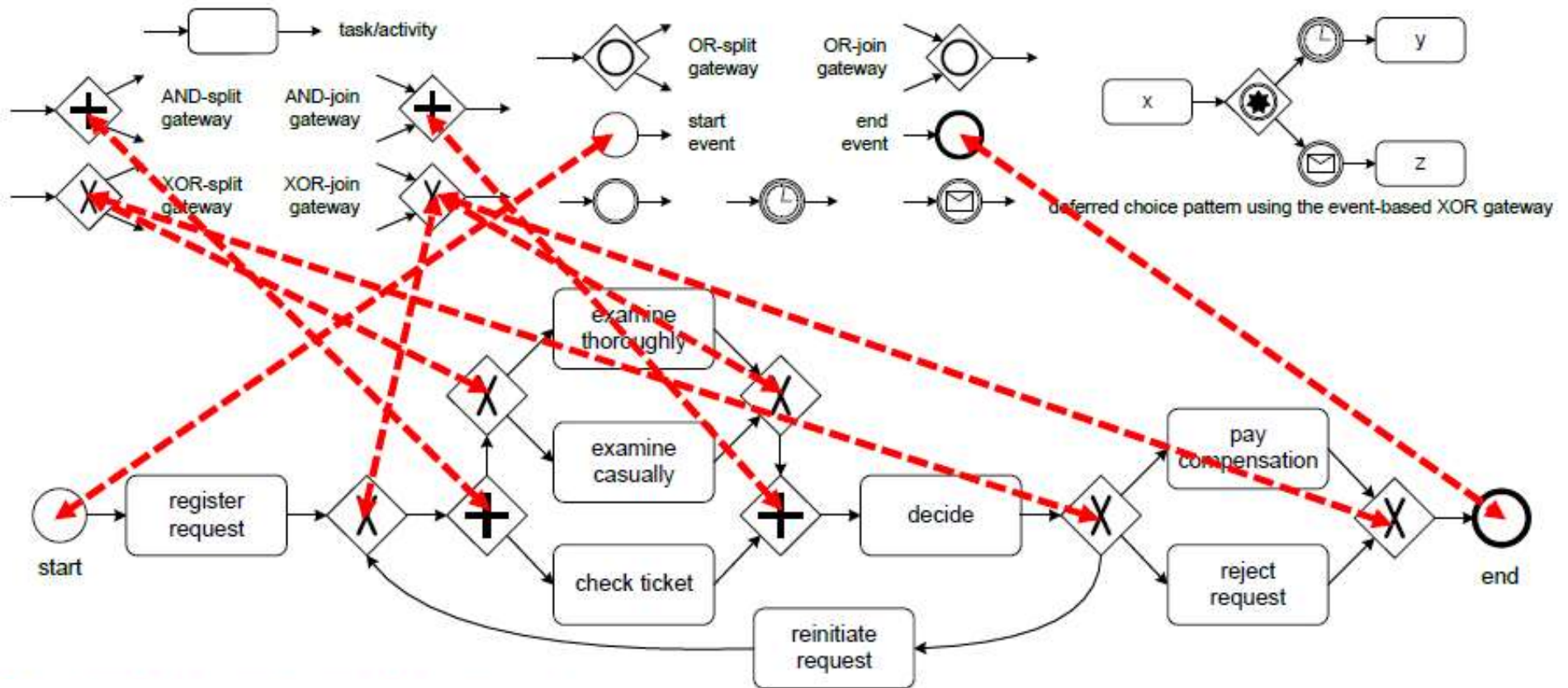


Case	Activity	Timestamp	Resource
432	register travel request (a)	18-3-2018:9.15	John
432	get support from local manager (b)	18-3-2018:9.25	Mary
432	check budget by finance (d)	18-3-2018:9.32	John
432	decide (e)	18-3-2018:9.36	Sue
432	accept request (g)	18-3-2018:9.48	Mary

A BPMN FONTOSABB ELEMEI:



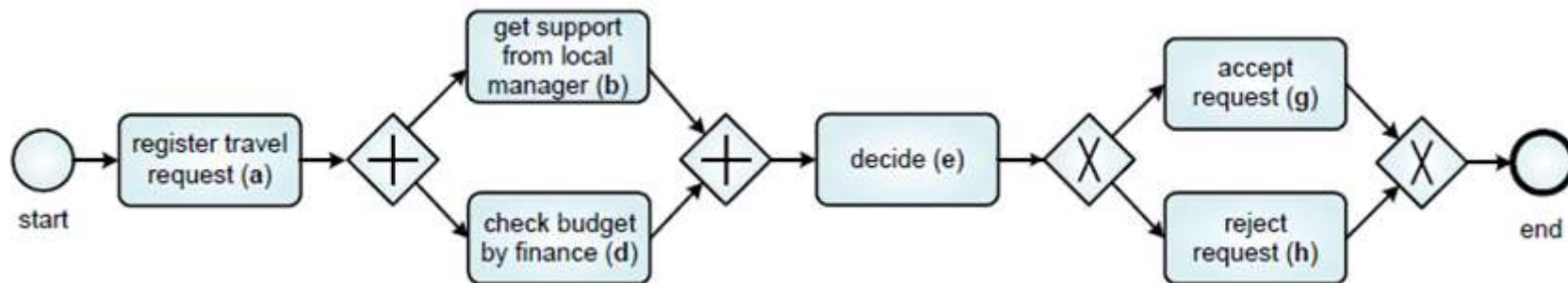
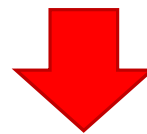
PÉLDA BPMN DIAGRAMRA



PLAY-IN: LOG → MODELL

EGYSZERŰ FOLYAMAT 4 FÉLE TRACE-HEZ

abdeg adbeg abdeg adbeh
abdeh abdeh adbeg abdeh
abdeh abdeh abdeh



REPLAY

acdeg

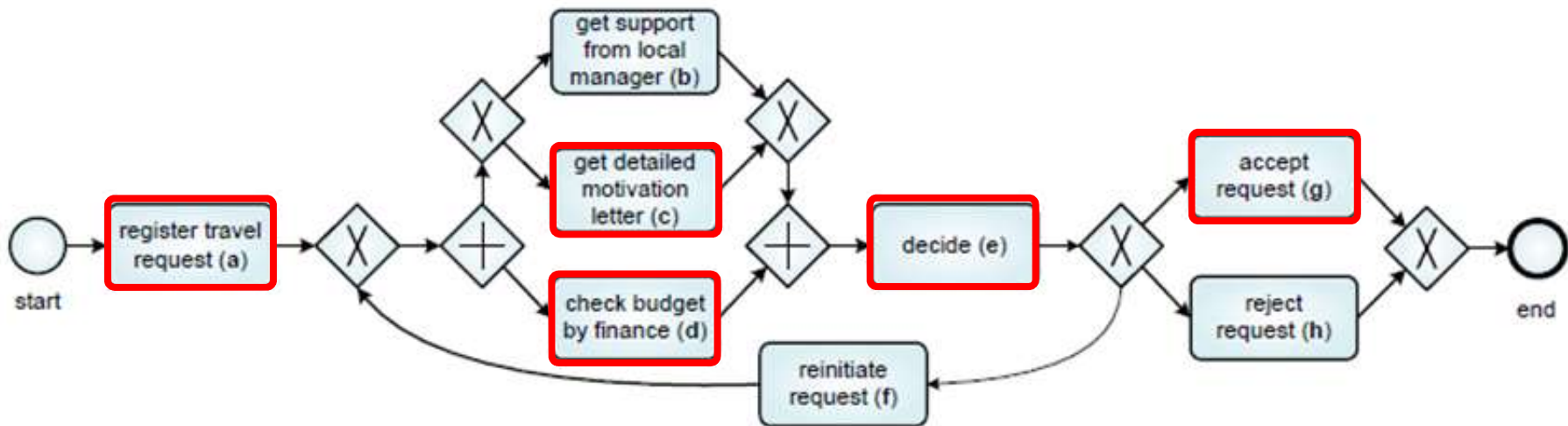


helyes

achdeg



helytelen

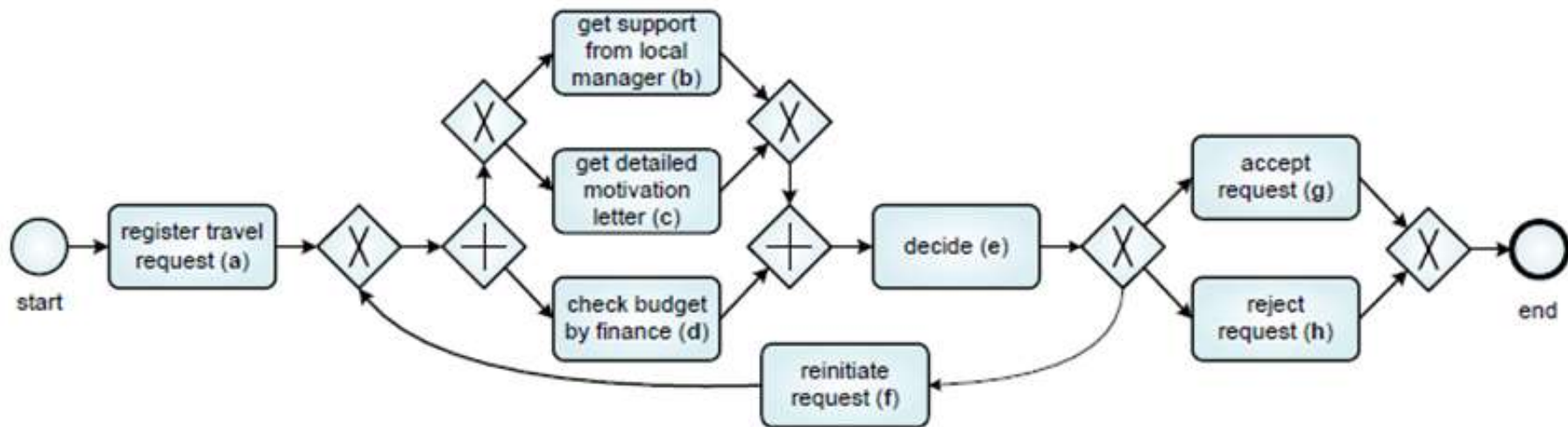


MELY TRACE-EK KÖVETHETŐK VÉGIG A MODELLBEN?

abdceh

aceh

adcefbdeg



MELY TRACE-EK KÖVETHETŐK VÉGIG A MODELLBEN?

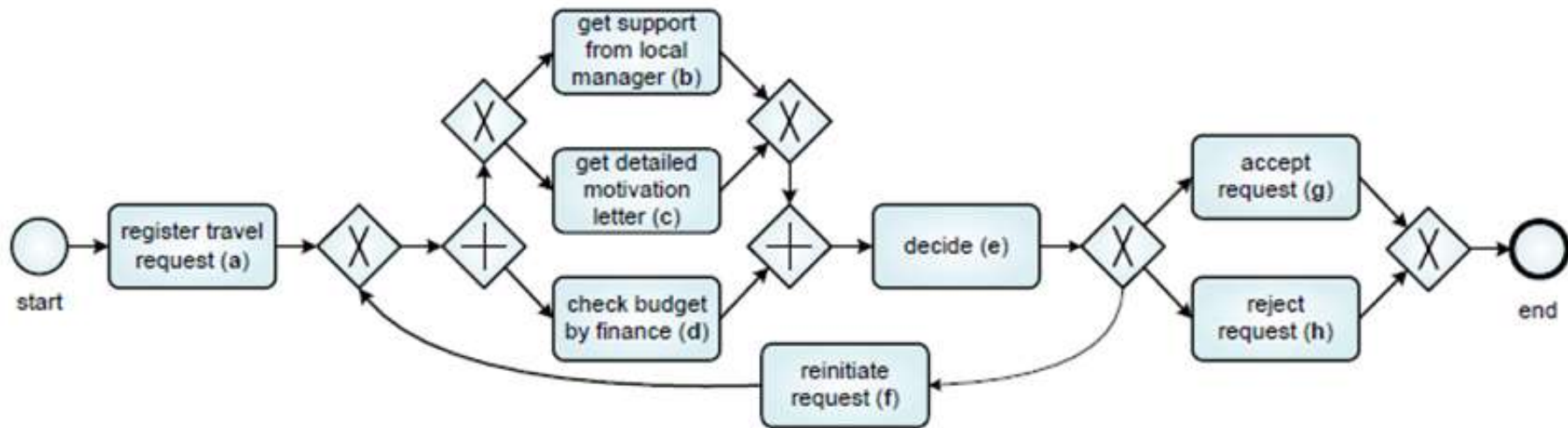
abdceh



aceh



adcefbdeg



WORKFLOW NETS

Gyakran dolgozunk munkafolyamat hálókkal (workflow net), amelyeknek jól meghatározott kezdetük és végük van.



A munkafolyamat hálók:

- anomáliáktól mentesek → megbízhatóság,
- a Petri hálók egy alosztályát jelentik,
- gyakran használatosak munkafolyamat-irányítási és üzleti folyamat menedzsment rendszereknél.

MI A SZEREPE A FOLYAMATI MODELLEKNEK?

Modellek szerepe:

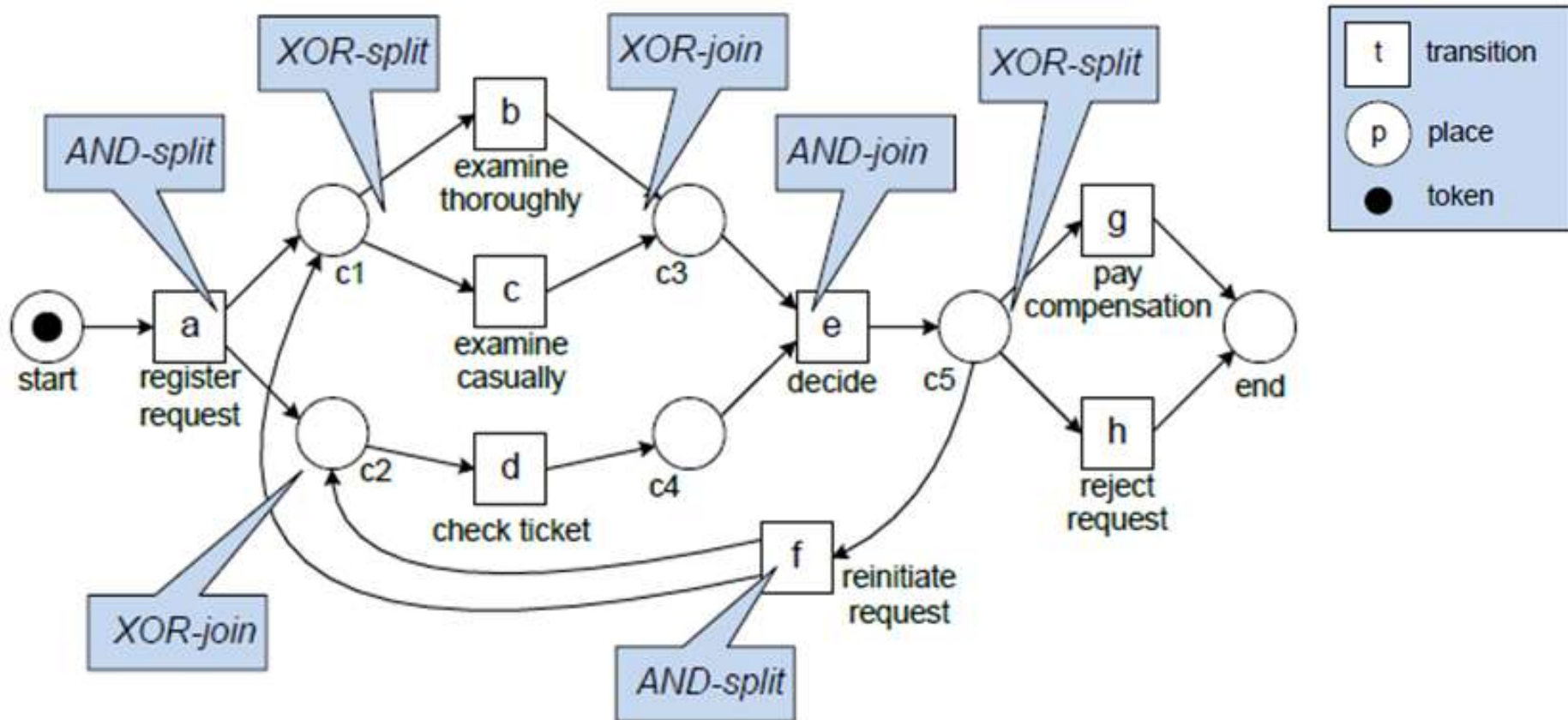
- következtetés a folyamatokról → újratervezés
- döntéseket hozunk a folyamatok közbülső részeiről → tervezés és ellenőrzés

Modellek használhatók:

- egyeztetni lehet a kötelezettségeket, elvárásokat,
- elemezni lehet az elvárásoknak való megfelelést,
- megjósolható a teljesítmény szimulációt használva,
- konfigurálhatunk egy workflow vagy BPM rendszert.

A KÜLÖNBÖZŐ MEGKÖZELÍTÉSEK KÜLÖNBÖZŐ JELÖLÉSEKET HASZNÁLNAK

Petri hálók



A MODELL ALAPÚ ELEMZÉS

A modell alapú elemzés főbb **típusai**:

- **Verifikáció** (hasonló a megbízhatóság ellenőrzéssel)
- **Teljesítmény elemzés** (szimuláció)

A modell alapú elemzés **korlátai**:

- Az ellenőrzés és teljesítmény elemzés attól függ, hogy milyen magas minőségű modell áll rendelkezésre.
- **Ha a modell és a valós lefutások között kicsi a hasonlóság, az elemzésnek nem sok értelme van.**
- Előfordul, hogy kicsi az összhang a kézzel készített modellek és a valóság között.

A folyamat bányászat célja ezen problémák kezelése, azáltal, hogy **közvetlen kapcsolatot valósít meg a modellek és az aktuális folyamati adatok között.**

HOGYAN TUDJUK EGY FOLYAMATNAK A MODELLJÉT ÚGY ELŐÁLLÍTANI, HOGY A LEGJOBBAN LEFEDJE A TRACE-EKET? PÉLDA: EGY CÉG TELEFONJAVÍTÁSI FOLYAMATA

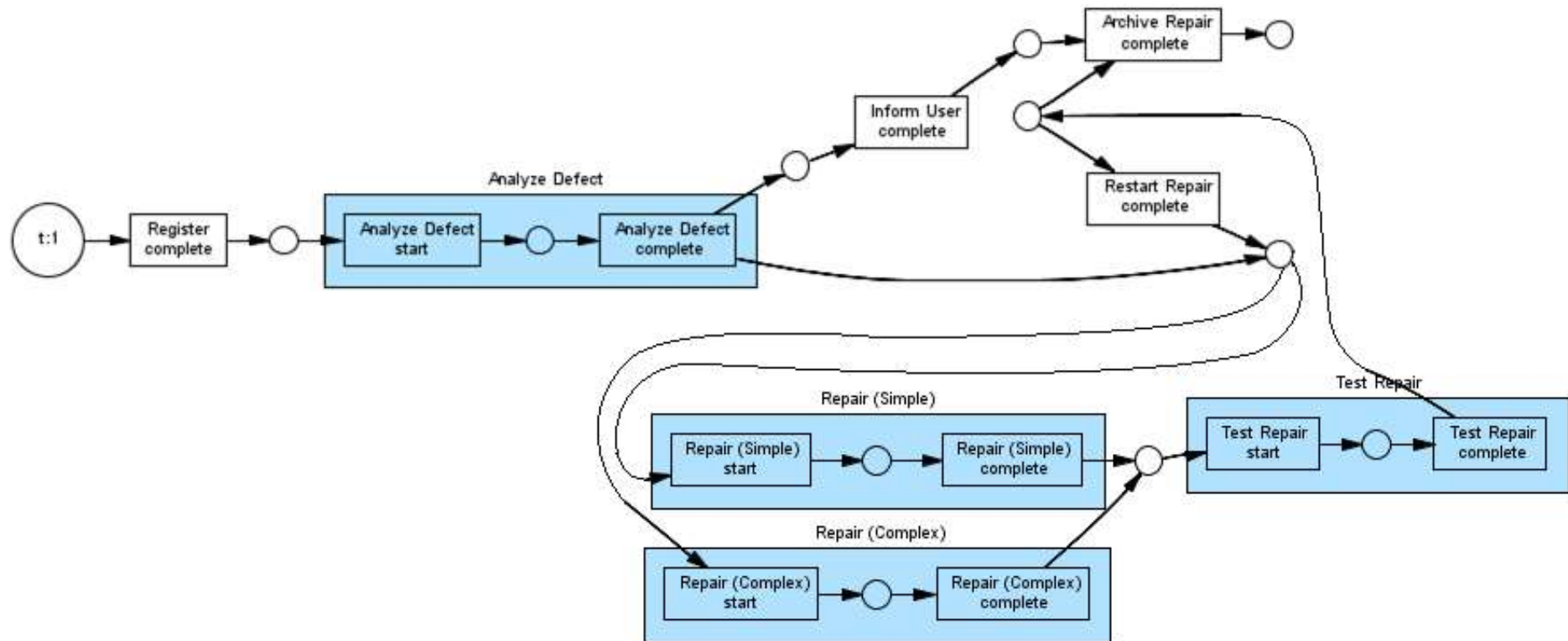
1. Három féle telefontípust képesek helyrehozni (T1, T2, T3).
2. A folyamat azzal indul, hogy regisztrálják a telefonkészüléket, amelyet a megbízó küldött.
3. Miután ez végbement, a telefon a **Probléma Detektálási (PD) részlegre** kerül. Itt elemzik a hiba okát és az általuk 10 féle javítási kategóriába sorolják.
4. Ha ez megtörtént, a készülék a hiba adataival együtt a **Javítási (R) részlegre** kerül.
5. A Javítási részlegnek két csoportja van.
Az egyik, amely az egyszerűbb hibákat javítja, a másik a bonyolultabb meghibásodások javításaiért felelős.
Léteznek olyan hibák is, amelyeket mindegyik csapat meg tud javítani.
6. Miután helyrehozták a készüléket a **Minőség Ellenőrzés** veszi át az irányítást, ahol megbizonyosodnak a hiba tényleges kijavításáról.
7. Ha a telefon újabb javításokra szorul, visszakerül a Javítási részlegre.
8. Ha a készülék immár hibátlan, az ügyet archiválják, és a telefont visszaküldik a megbízónak.
9. Ha bizonyos javítás után sem lehet helyrehozni a telefont, az ügyet akkor is lezárják, és az ügyfél kap egy teljesen új készüléket.


```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- MXML version 1.0 -->
<WorkflowLog xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="http://is.tm.tue.nl/research/processmining/WorkflowLog.xsd"
  description="CPN Tools simulation log">
  <Source program="CPN Tools simulation"/>
  <Process id="DEFAULT" description="Simulated process">
    <ProcessInstance id="1" description="Simulated process instance">
      <AuditTrailEntry>
        <WorkflowModelElement>Register</WorkflowModelElement>
        <EventType >complete</EventType>
        <Timestamp>1970-01-02T12:23:00.000+01:00</Timestamp>
        <Originator>System</Originator>
      </AuditTrailEntry>
      <AuditTrailEntry>
        <WorkflowModelElement>Analyze Defect</WorkflowModelElement>
        <EventType >start</EventType>
        <Timestamp>1970-01-02T12:23:00.000+01:00</Timestamp>
        <Originator>Tester3</Originator>
      </AuditTrailEntry>
      <AuditTrailEntry>
        <Data>
          <Attribute name = "phoneType">T2 </Attribute>
          <Attribute name = "defectType">6 </Attribute>
        </Data>
        <WorkflowModelElement>Analyze Defect</WorkflowModelElement>
        <EventType >complete</EventType>
        <Timestamp>1970-01-02T12:30:00.000+01:00</Timestamp>
        <Originator>Tester3</Originator>
      </AuditTrailEntry>
    </ProcessInstance>
  </Process>
</WorkflowLog>
```

...

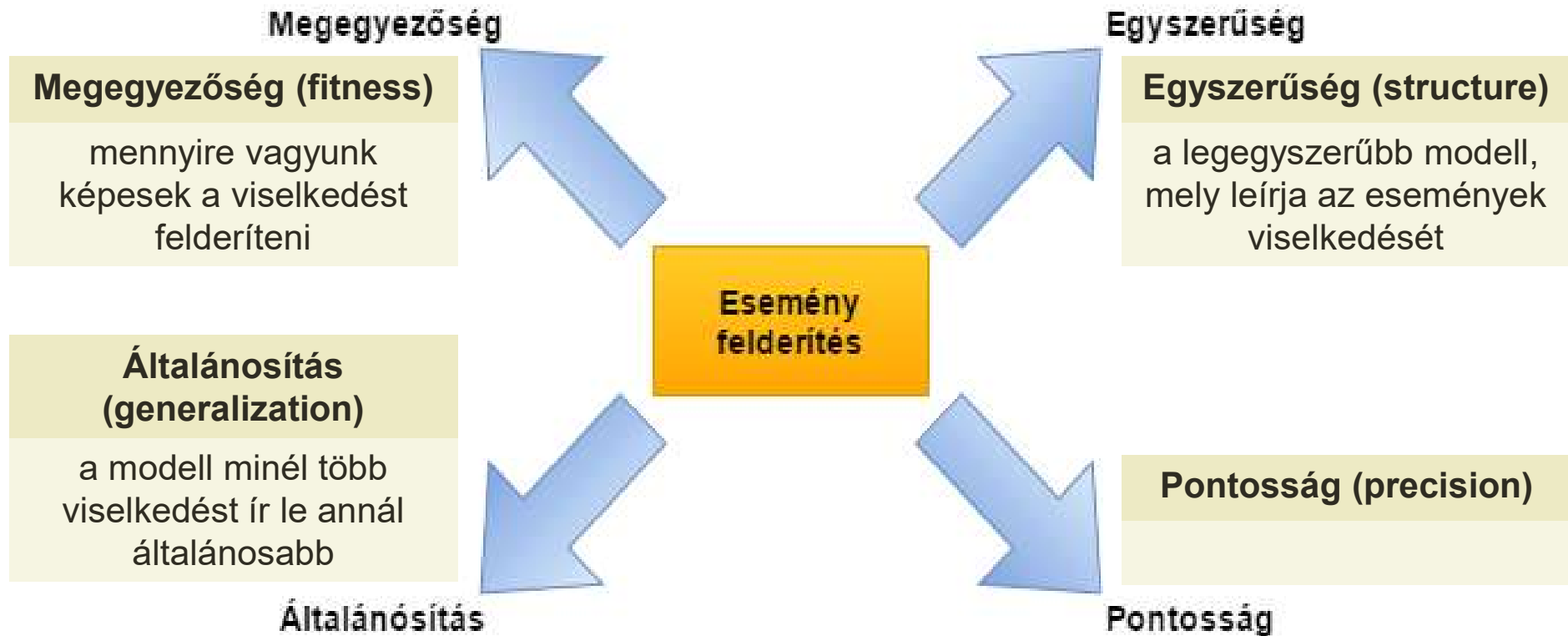
A FOLYAMAT
ESEMÉNY-
NAPLÓJÁNAK
EGY
RÉSZLETE

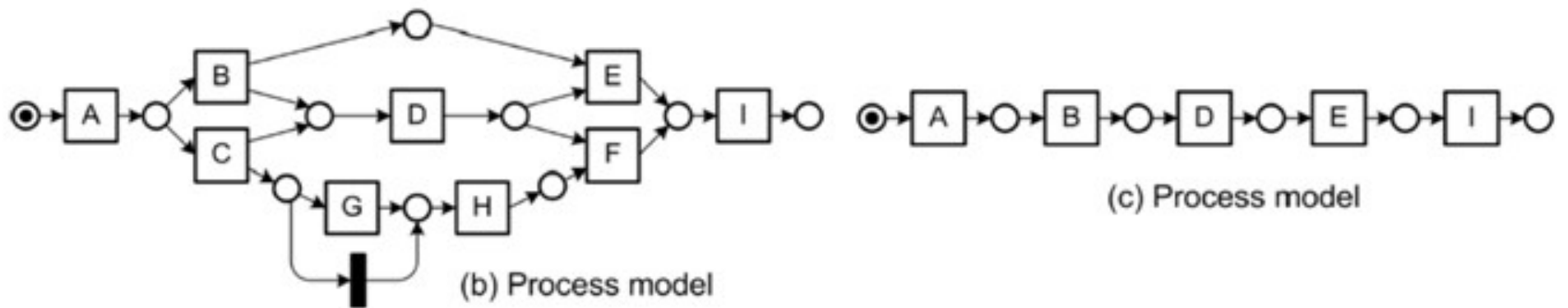
A FELADAT PETRI HÁLÓJA – ALPHA ALGORITMUS HASZNÁLATA



NÉGY FŐ MODELL TULAJDONSÁG

4 egymással versengő jellemzőt kell figyelembe venni:





fitness +
precision +
generalization +
structure +

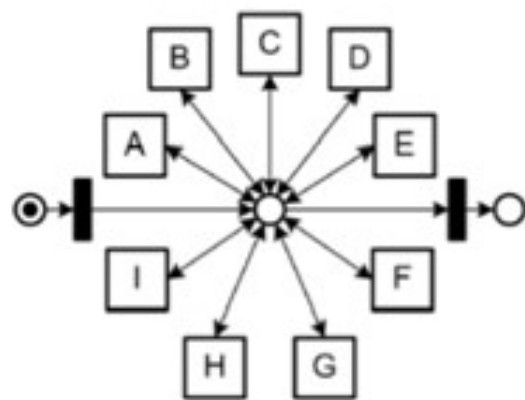
fitness +
precision -
generalization +
structure +

No. of Instances	Log Traces
1207	ABDEI
145	ACDGHFI
56	ACGDHFI
23	ACHDFI
28	ACDHF

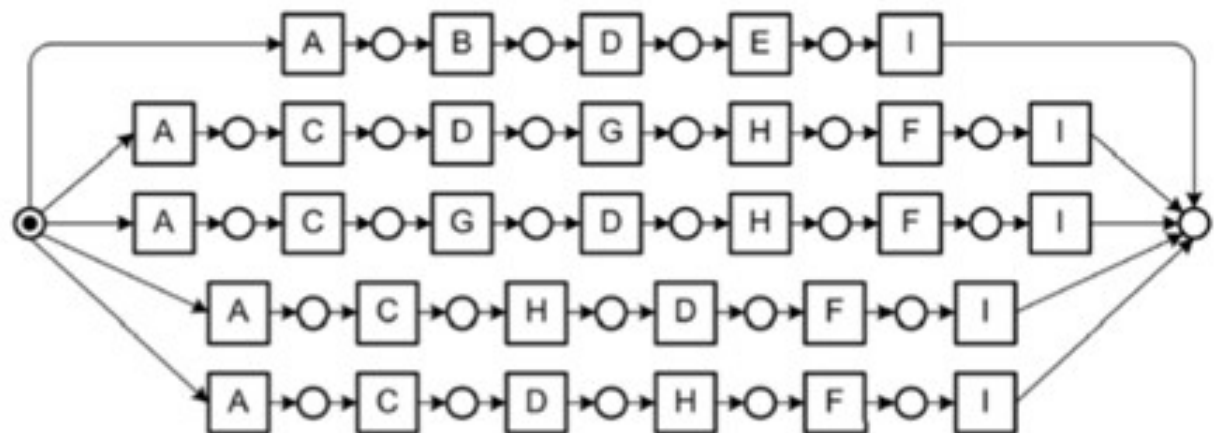
(a) Event Log

fitness -
precision +
generalization -
structure +

fitness +
precision +
generalization -
structure -

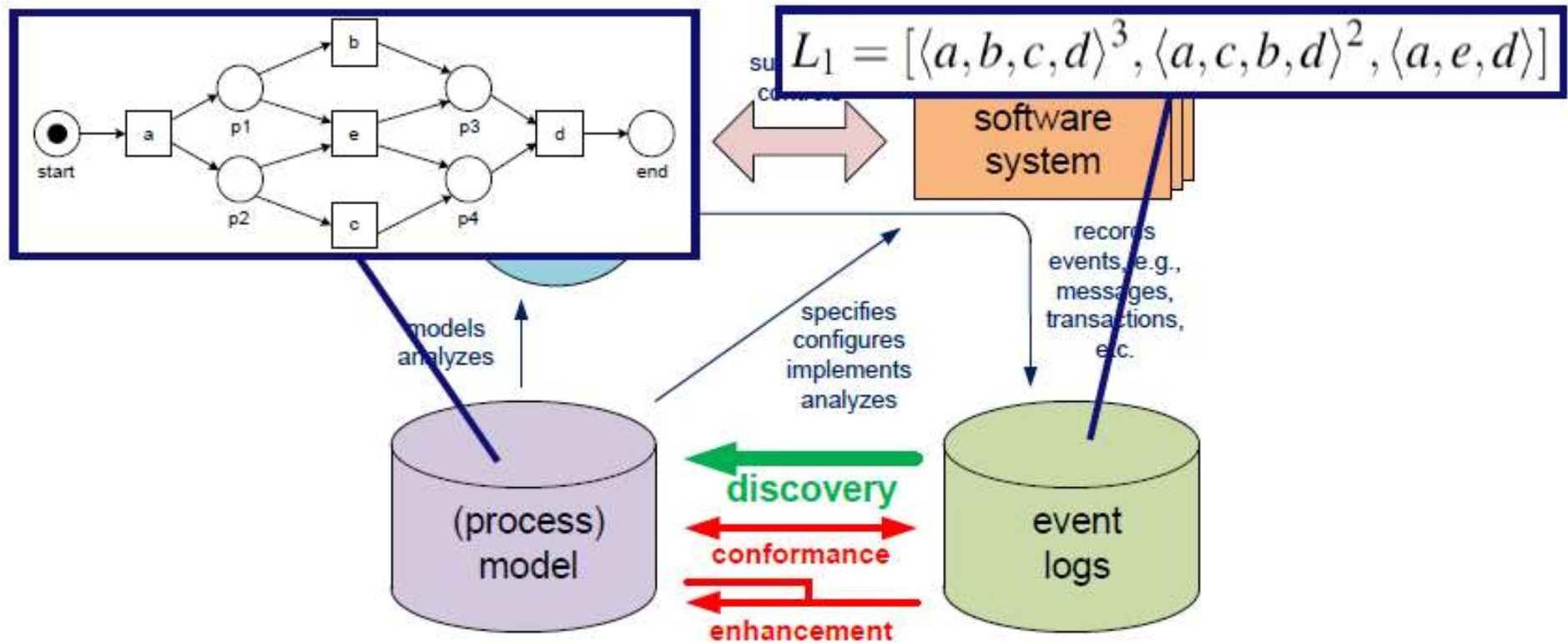


(d) Process model



(e) Process model

AZ ALFA ALGORITMUS HASZNÁLATA FOLYAMATI LOGOKBÓL MODELL (WF-NET) KINYERÉSÉHEZ (PLAY-IN)



NAPLÓFÁJL EGYSZERŰSÍTÉSE

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 Hop	iPhone5S	1
9902	check stock	22-1-2014@10.57	Pete Scott	iPhone5S	2

[<register_order, check_stock, ship_order, handle_payment>
<register_order, check_stock, cancel_order>
<register_order, check_stock> , ...]

Egyszerűsítsük a rendelkezésre álló naplófájlt, hogy egy WF-net előállítható legyen.
Csak az esetenkénti eseménylefutási sorrendek fontosak.

LOG, TRACE-EK

$$L = [\langle a, b, c, d \rangle^3, \langle a, c, b, d \rangle^2, \langle a, e, d \rangle]$$

Diagram illustrating the structure of a log L. The log is a set of traces, each represented as a sequence of activities (tevékenység) and their occurrence counts (trace előfordulási mennyisége a logban). The log L is shown as a set of three traces: $\langle a, b, c, d \rangle^3$, $\langle a, c, b, d \rangle^2$, and $\langle a, e, d \rangle$. The first trace is labeled "trace" and the entire set is labeled "log".

Az esemény log **trace**-eknek a halmaza (ugyanaz a trace többször is megjelenhet a logban).

A trace **tevékenység**ek neveinek sorozatát tartalmazza (minden egyéb tulajdonságtól most elvonatkoztatunk).

A cél, hogy a rendelkezésre álló trace-ekből felépítsünk egy olyan **modell**t, amelyre minden trace illeszkedik.

PÉLDA LOG

Minimális információ a logban:

- eset azonosító - **Case**
- feladat azonosító - **Task**

Szekvenciák:

- 1: ABCD
- 2: ACBD
- 3: ABCD
- 4: ACBD
- 5: EF



3 lehetséges szekvencia:

- ABCD
- ACBD
- EF

$$L = [\langle A, B, C, D \rangle^2, \langle A, C, B, D \rangle^2, \langle E, F \rangle]$$

Case	Task
1	A
2	A
3	A
3	B
1	B
1	C
2	C
4	A
2	B
2	D
5	E
4	C
1	D
3	C
3	D
4	B
5	F
4	D

>, →, ||, # KAPCSOLATOK

>	direkt sorrend	$x > y$	az x esetet közvetlenül az y eset követi
→	okozati viszony	$x \rightarrow y$	ha $x > y$ és nem $y > x$
	párhuzamosság	$x y$	ha $x > y$ és $y > x$
#	választás	$x \# y$	ha nem $x > y$ és nem $y > x$

Case	Task
1	A
2	A
3	A
3	B
1	B
1	C
2	C
4	A
2	B
2	D
5	E
4	C
1	D
3	C
3	D
4	B
5	F
4	D



ABCD
ACBD
EF



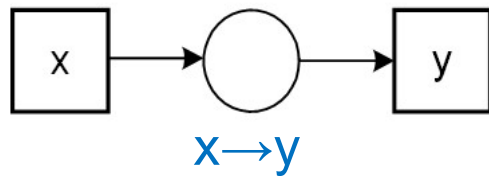
A > B
A > C
B > C
B > D
C > B
C > D
E > F

A → B
A → C
B → D
C → D
E → F

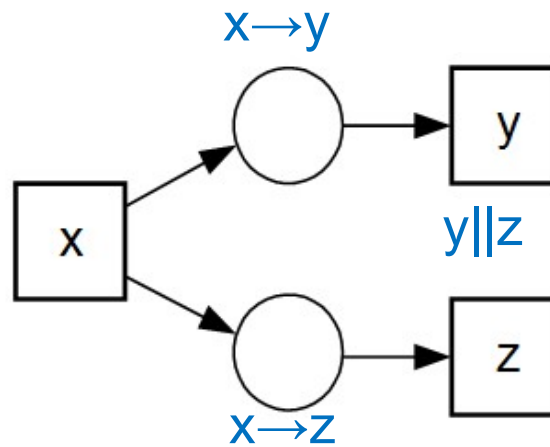
B || C
C || B

ALAP ÖTLETEK

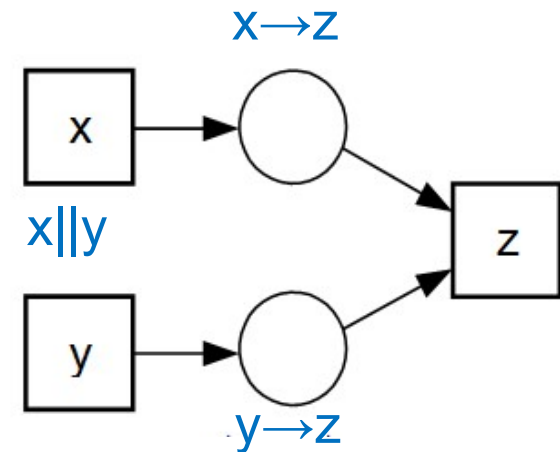
Szekvencia minta



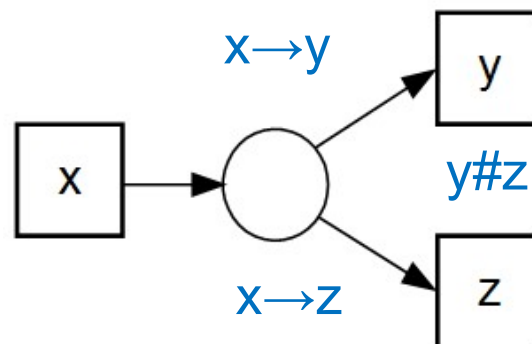
AND-split minta



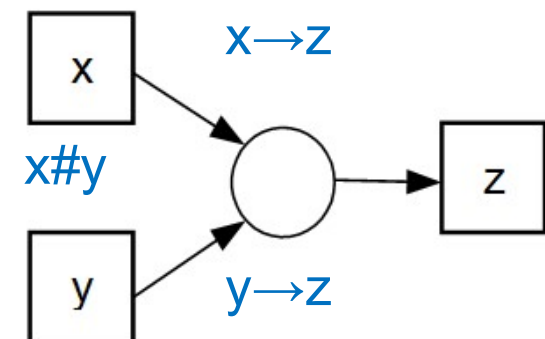
AND-join minta



XOR-split minta



XOR-join minta



AZ ALFA ALGORITMUS

Legyen W munkafolyamat napló T műveletekkel. $\alpha(W)$ a következő képen van definiálva:

1. $T_W = \{ t \in T \mid \exists \sigma \in W \ t \in \sigma \}$ műveletek halmaza
2. $T_I = \{ t \in T \mid \exists \sigma \in W \ t = \text{first}(\sigma) \}$ kezdő műveletek halmaza
3. $T_O = \{ t \in T \mid \exists \sigma \in W \ t = \text{last}(\sigma) \}$ befejező műveletek halmaza
4. $X_W = \{ (A, B) \mid A \subseteq T_W \wedge A \neq \emptyset \wedge B \subseteq T_W \wedge B \neq \emptyset \wedge \forall a \in A \forall b \in B \ a \rightarrow_W b \wedge \forall a_1, a_2 \in A \ a_1 \#_W a_2 \wedge \forall b_1, b_2 \in B \ b_1 \#_W b_2 \}$ Vesszük a T_I ké nem üres részalmazát
5. $Y_W = \{ (A, B) \in X \mid \forall (A', B') \in X \ A \subseteq A' \wedge B \subseteq B' \Rightarrow (A, B) = (A', B') \}$ Töröljük a nem maximális párokat
6. $P_W = \{ p_{(A,B)} \mid (A, B) \in Y_W \} \cup \{ i_W, o_W \}$ Definiáljuk a helyek halmazát
7. $F_W = \{ (a, p_{(A,B)}) \mid (A, B) \in Y_W \wedge a \in A \} \cup \{ (p_{(A,B)}, b) \mid (A, B) \in Y_W \wedge b \in B \} \cup \{ (i_W, t) \mid t \in T_I \} \cup \{ (t, o_W) \mid t \in T_O \}$ Meghatározzuk a folyamatbeli relációkat
8. $\alpha(W) = (P_W, T_W, F_W)$ az Alpha algoritmussal előállított modell

A PÉLDA ÁBRÁZOLÁSA

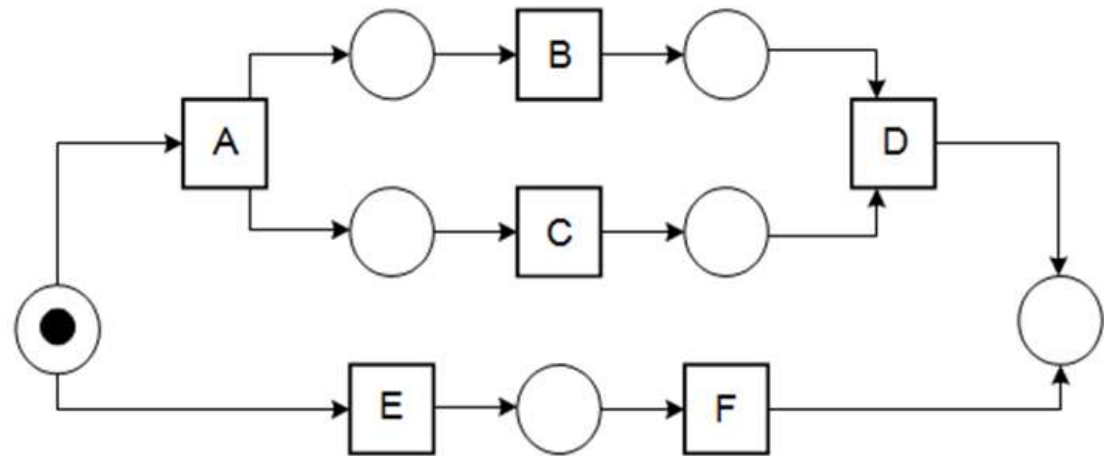
Log (W):

Case	Task
1	A
2	A
3	A
3	B
1	B
1	C
2	C
4	A
2	B
2	D
5	E
4	C
1	D
3	C
3	D
4	B
5	F
4	D

???



Kimenet: $\alpha(W)$



Kapcsolatok:

A > B	A → B
A > C	A → C
B > C	B → D
B > D	C → D
C > B	E → F
C > D	B C
E > F	C B

$$L = [\langle a, b, c, d \rangle^2, \langle a, c, b, d \rangle^2, \langle e, f \rangle]$$

LÁBNYOM TÁBLÁZAT

ABCD
ACBD
EF



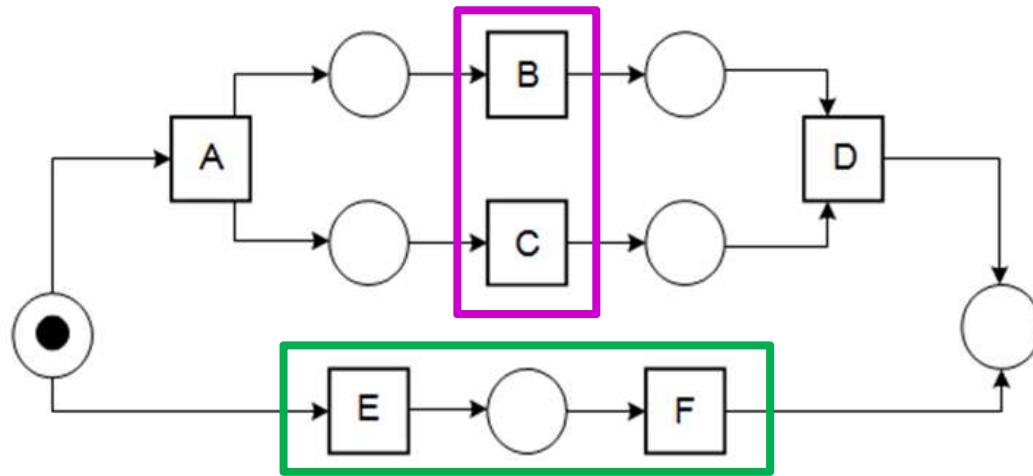
A→B
A→C
B→D
C→D
E→F

B || C
C || B



	A	B	C	D	E	F
A	#	→	→	#	#	#
B	←	#		→	#	#
C	←		#	→	#	#
D	#	←	←	#	#	#
E	#	#	#	#	#	→
F	#	#	#	#	←	#

A FELFEDEZETT MODELLNEK AZONOS A TARTALMA MINT A LÁBNYOM TÁBLÁZATNAK



ABCD
ACBD
EF



A → B
A → C
B → D
C → D
E → F

B || C
C || B



	A	B	C	D	E	F
A	#	→	→	#	#	#
B	←	#		→	#	#
C	←		#	→	#	#
D	#	←	←	#	#	#
E	#	#	#	#	#	→
F	#	#	#	#	←	#

AZ ALFA ALGORITMUS ÉRTELMEZÉSE PÉLDÁN KERESZTÜL

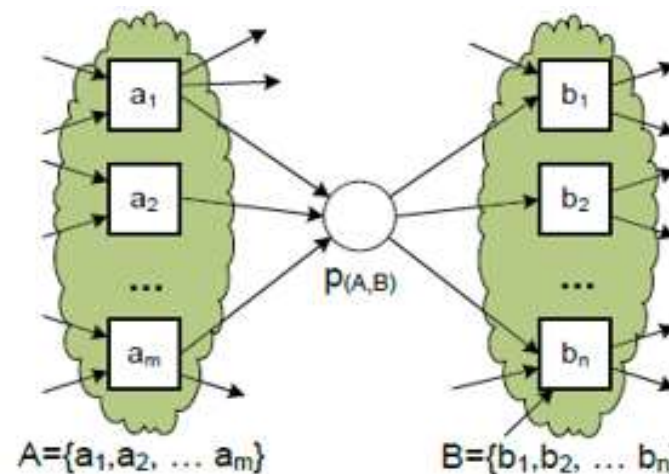
$$L = [\langle a, b, c, d \rangle^3, \langle a, c, b, d \rangle^2, \langle a, e, d \rangle]$$

	a	b	c	d	e
a	#	→	→	#	→
b	←	#		→	#
c	←		#	→	#
d	#	←	←	#	←
e	←	#	#	→	#

1. $T_L = \{t \in T \mid \exists \sigma \in L t \in \sigma\}$, **műveletek**
 $T_L = \{a, b, c, d, e\}$

2. $T_I = \{t \in T \mid \exists \sigma \in L t = \text{first}(\sigma)\}$ **kezdő művelet(ek)**
 $T_I = \{a\}$

3. $T_O = \{t \in T \mid \exists \sigma \in L t = \text{last}(\sigma)\}$ **befejező művelet(ek)**
 $T_O = \{d\}$



AZ ALFA ALGORITMUS ÉRTELMEZÉSE PÉLDÁN KERESZTÜL

(A,B) párok meghatározása

A és B halmaz a log-beli műveletek
részhalmaza

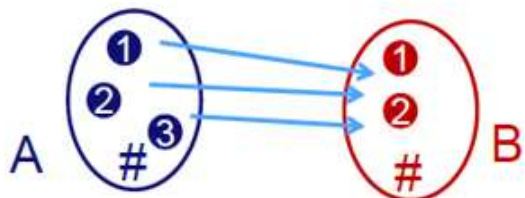
$$4. X_L = \{ (A,B) \mid A \subseteq T_L \wedge A \neq \emptyset \wedge B \subseteq T_L \wedge B \neq \emptyset$$

$$\wedge \forall a \in A \forall b \in B a \rightarrow_L b \quad \text{A-t követi B}$$

$$\wedge \forall a_1, a_2 \in A a_1 \#_L a_2 \quad \text{A halmazból választások}$$

$$\wedge \forall b_1, b_2 \in B b_1 \#_L b_2 \} \quad \text{B halmazból választások}$$

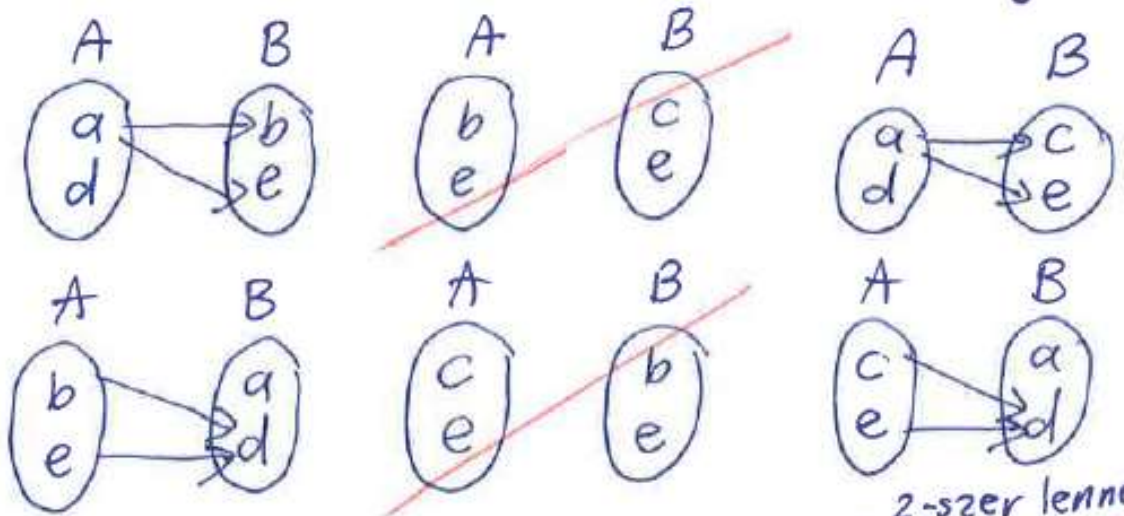
$$X_L = \{ (\{a\}, \{b\}), (\{a\}, \{c\}), (\{a\}, \{e\}), (\{b\}, \{d\}), (\{c\}, \{d\}), (\{e\}, \{d\}), \\ (\{a\}, \{b, e\}), (\{a\}, \{c, e\}), (\{b, e\}, \{d\}), (\{c, e\}, \{d\}) \}$$



	a_1	a_2	...	a_m	b_1	b_2	...	b_n
a_1	#	#	...	#	→	→	...	→
a_2	#	#	...	#	→	→	...	→
...
a_m	#	#	...	#	→	→	...	→
b_1	←	←	...	←	#	#	...	#
b_2	←	←	...	←	#	#	...	#
...
b_n	←	←	...	←	#	#	...	#

$$L = [\langle a, b, c, d \rangle^3, \langle a, c, b, d \rangle^2, \langle a, e, d \rangle]$$

$a \rightarrow b$ $a \# d$
 $a \rightarrow c$ $b \# e$
 $a \rightarrow e$ $c \# e$
 $c \rightarrow d$
 $b \rightarrow d$ $b \parallel c$
 $e \rightarrow d$



$$X_L = \{ \cancel{(\{a\}, \{b\})}; \cancel{(\{a\}, \{e\})}; (\{a\}, \{b, e\}); \cancel{(\{a\}, \{e\})}; \cancel{(\{b\}, \{d\})}; \cancel{(\{c\}, \{d\})}; \cancel{(\{e\}, \{d\})}; (\{c, e\}, \{d\}); (\{b, e\}, \{d\}); \cancel{(\{a\}, \{c\})}; (\{a\}, \{c, e\}) \}$$

AZ ALFA ALGORITMUS ÉRTELMEZÉSE PÉLDÁN KERESZTÜL

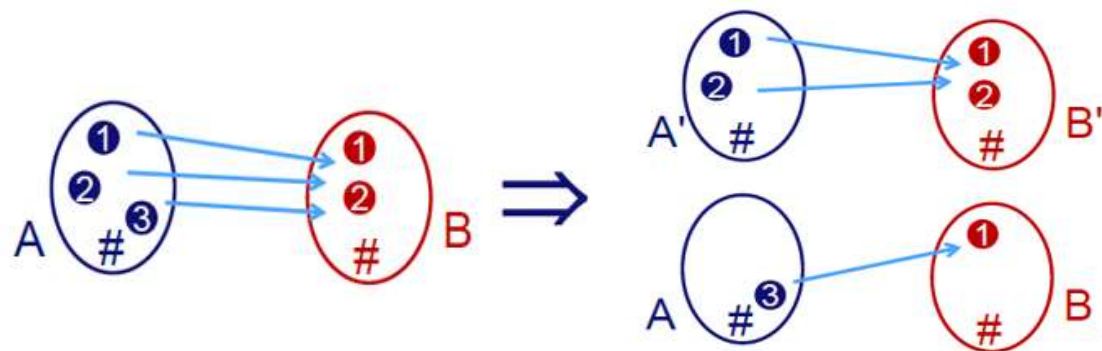
nem maximális (A,B) párok törlése

$$5. Y_L = \{ (A,B) \in X_L \mid \forall_{(A',B') \in X_L} A \subseteq A' \wedge B \subseteq B' \Rightarrow (A,B) = (A',B') \}$$

$$X_L = \{ (\{a\}, \{b\}), (\{a\}, \{c\}), (\{a\}, \{e\}), (\{b\}, \{d\}), (\{c\}, \{d\}), (\{e\}, \{d\}), \\ (\{a\}, \{b, e\}), (\{a\}, \{c, e\}), (\{b, e\}, \{d\}), (\{c, e\}, \{d\}) \}$$



$$Y_L = \{ (\{a\}, \{b, e\}), (\{a\}, \{c, e\}), (\{b, e\}, \{d\}), (\{c, e\}, \{d\}) \}$$

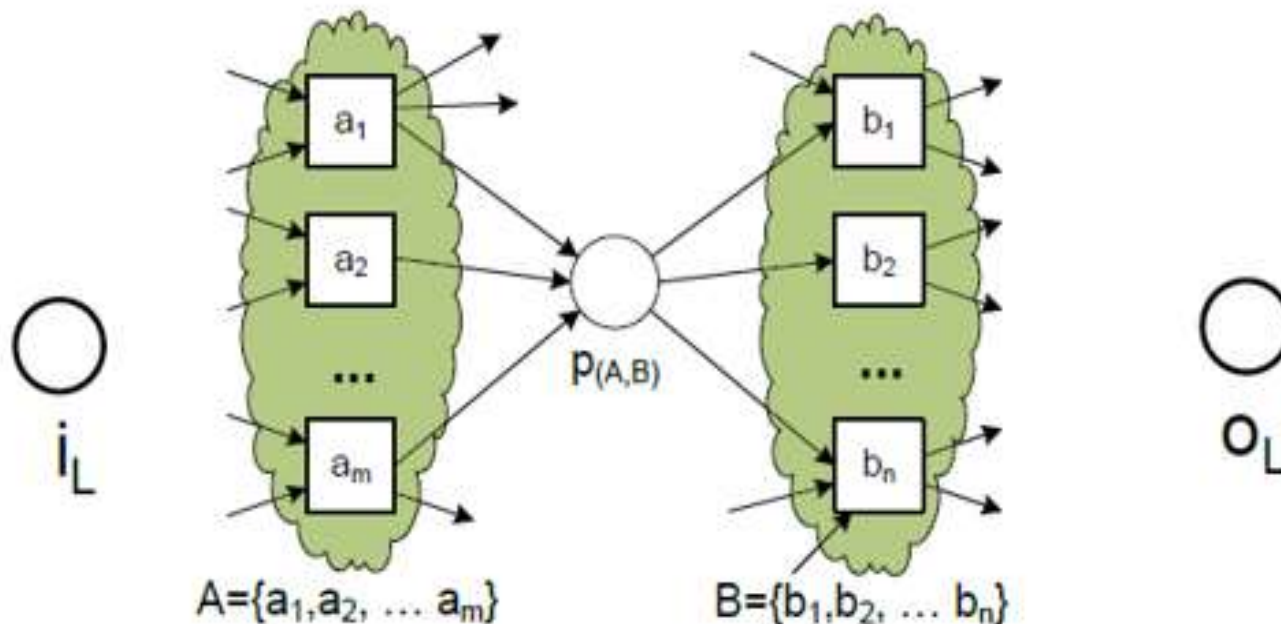


AZ ALFA ALGORITMUS ÉRTELMEZÉSE PÉLDÁN KERESZTÜL

$p(A,B)$ helyek meghatározása (A,B) párokból

$$6. P_L = \{ p_{(A,B)} \mid (A,B) \in Y_L \} \cup \{ i_L, o_L \}$$

$$P_L = \{ P_{(\{a\},\{b,e\})}, P_{(\{a\},\{c,e\})}, P_{(\{b,e\},\{d\})}, P_{(\{c,e\},\{d\})} \} \cup \{ i_L, o_L \}$$



AZ ALFA ALGORITMUS ÉRTELMEZÉSE PÉLDÁN KERESZTÜL

folyamatbeli **relációk** meghatározása
(milyen helyek és műveletek közt lesz nyíl)

$$7. F_L = \{ (a, p_{(A,B)}) \mid (A,B) \in Y_L \wedge a \in A \} \\ \cup \{ (p_{(A,B)}, b) \mid (A,B) \in Y_L \wedge b \in B \} \\ \cup \{ (i_L, t) \mid t \in T_I \} \cup \{ (t, o_L) \mid t \in T_O \}$$

$$F_L = \left\{ \begin{array}{l} (i_L, a), (d, o_L), \\ (a, P_{(\{a\},\{b,e\})}), (a, P_{(\{a\},\{c,e\})}), (P_{(\{a\},\{b,e\})}, b), (P_{(\{a\},\{b,e\})}, e), \\ (P_{(\{a\},\{c,e\})}, c), (P_{(\{a\},\{c,e\})}, e), (b, P_{(\{b,e\},\{d\})}), (e, P_{(\{b,e\},\{d\})}), \\ (c, P_{(\{c,e\},\{d\})}), (e, P_{(\{c,e\},\{d\})}), (P_{(\{b,e\},\{d\})}, d), (P_{(\{c,e\},\{d\})}, d) \end{array} \right\}$$

AZ ALFA ALGORITMUS ÉRTELMEZÉSE PÉLDÁN KERESZTÜL

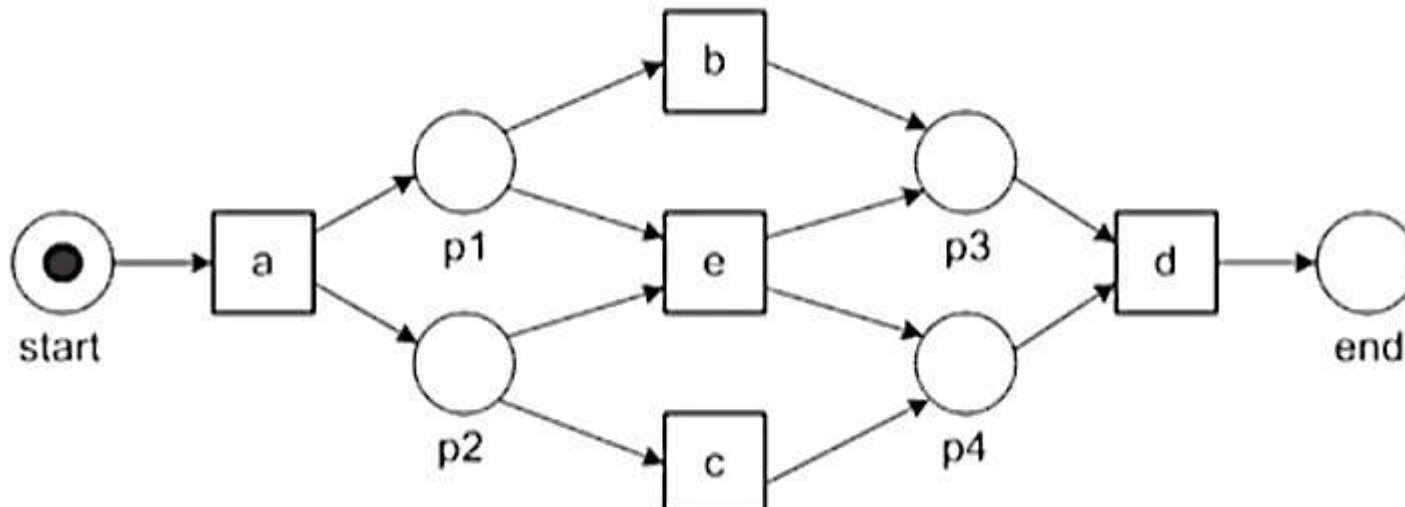
az algoritmus kimenete: munkafolyamat háló (helyek, műveletek, relációk)

$$8. \alpha(L) = (P_L, T_L, F_L)$$

$$P_L = \{P_{(\{a\},\{b,e\})}, P_{(\{a\},\{c,e\})}, P_{(\{b,e\},\{d\})}, P_{(\{c,e\},\{d\})}\} \cup \{i_L, o_L\}$$

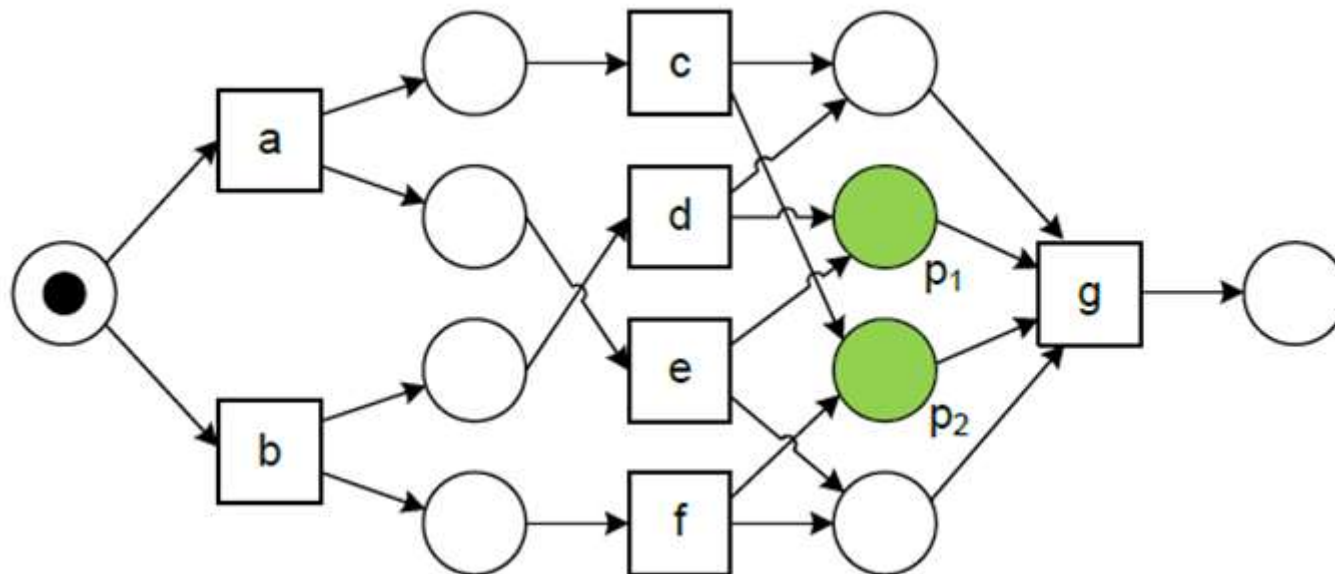
$$T_L = \{a, b, c, d, e\}$$

$$F_L = \left\{ \begin{array}{l} (i_L, a), (d, o_L), \\ (a, P_{(\{a\},\{b,e\})}), (a, P_{(\{a\},\{c,e\})}), (P_{(\{a\},\{b,e\})}, b), (P_{(\{a\},\{b,e\})}, e), \\ (P_{(\{a\},\{c,e\})}, c), (P_{(\{a\},\{c,e\})}, e), (b, P_{(\{b,e\},\{d\})}), (e, P_{(\{b,e\},\{d\})}), \\ (c, P_{(\{c,e\},\{d\})}), (e, P_{(\{c,e\},\{d\})}), (P_{(\{b,e\},\{d\})}, d), (P_{(\{c,e\},\{d\})}, d) \end{array} \right\}$$



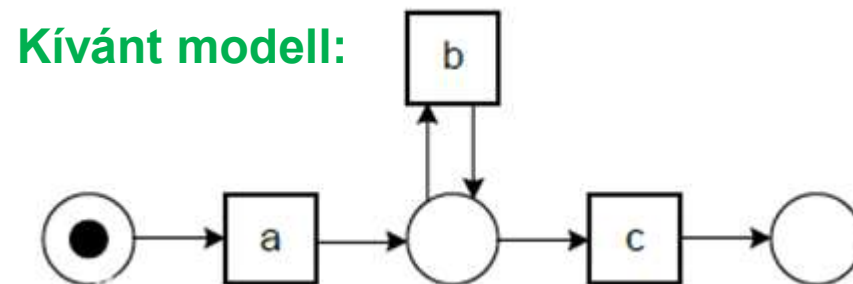
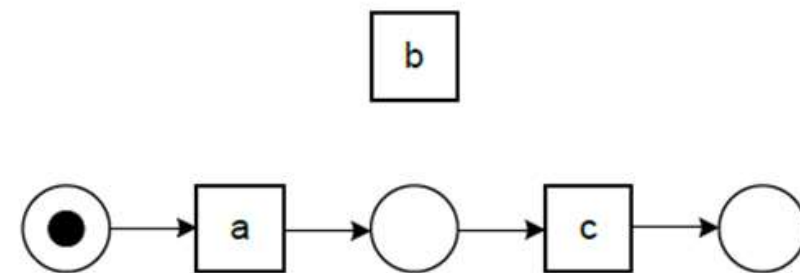
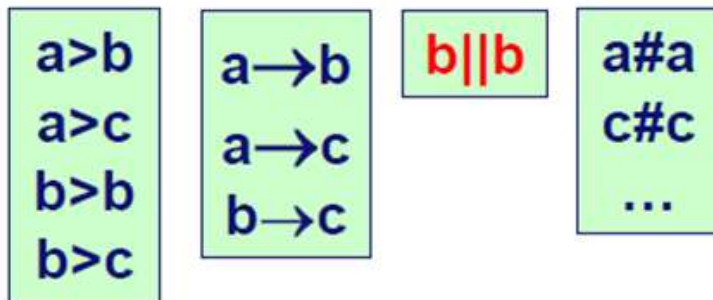
HALLGATÓLAGOS HELYEK A MODELLBEN: P1 ÉS P2

$$L = [\langle a, c, e, g \rangle^2, \langle a, e, c, g \rangle^3, \langle b, d, f, g \rangle^2, \langle b, f, d, g \rangle^4]$$



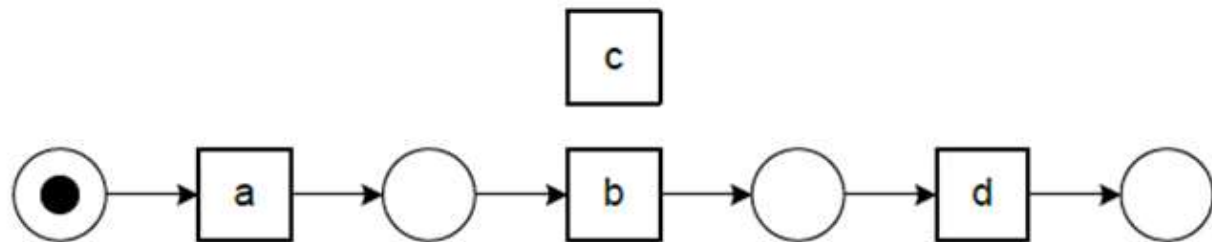
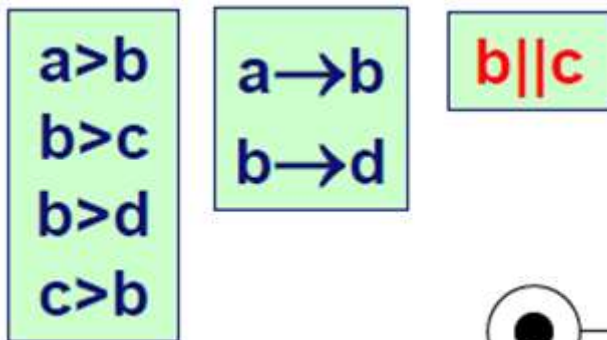
HUOK A MODELLBEN

$$L = [\langle a, c \rangle^2, \langle a, b, c \rangle^3, \langle a, b, b, c \rangle^2, \langle a, b, b, b, b, c \rangle^1]$$

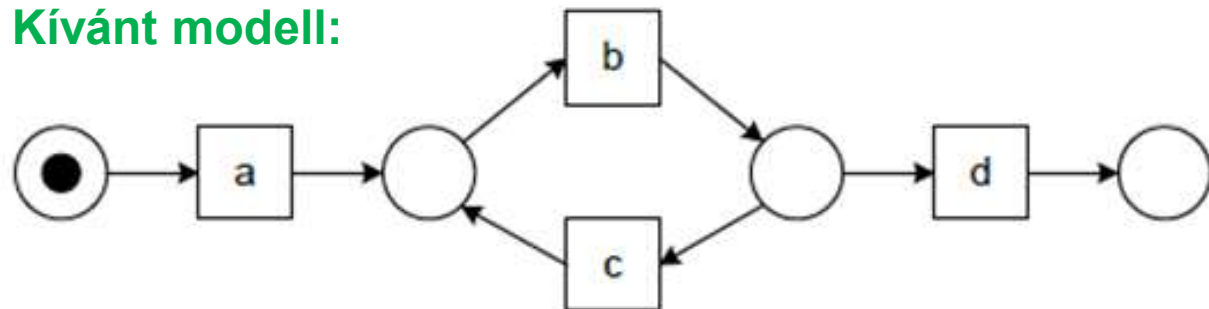


HUOK A MODELLBEN

$$L = [\langle a, b, d \rangle^3, \langle a, b, c, b, d \rangle^2, \langle a, b, c, b, c, b, d \rangle]$$

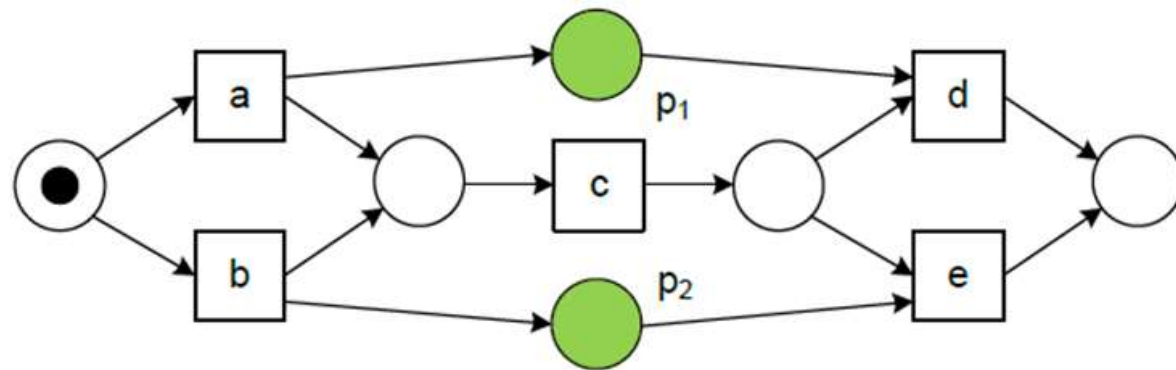
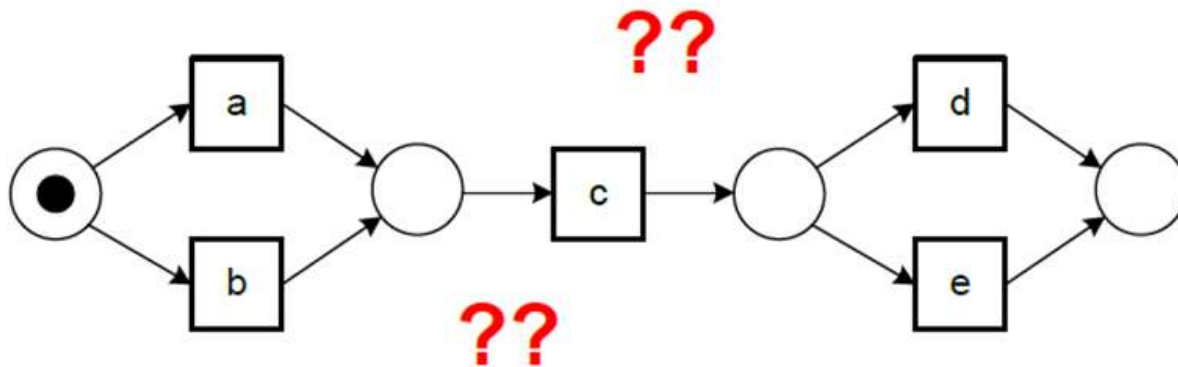


Kívánt modell:



TOVÁBBI PROBLÉMÁK

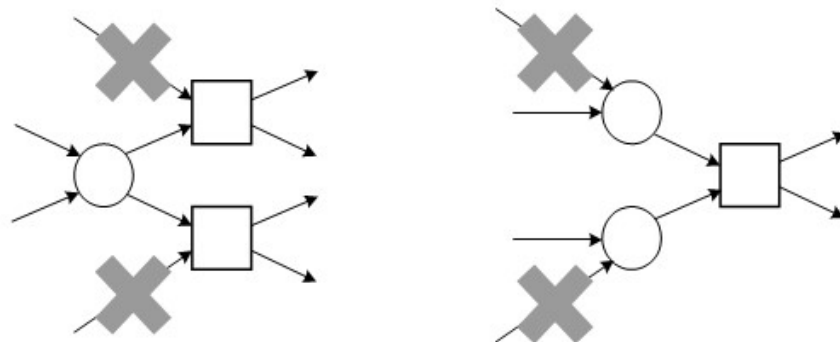
$$L = [\langle a, c, d \rangle^{45}, \langle b, c, e \rangle^{42}]$$



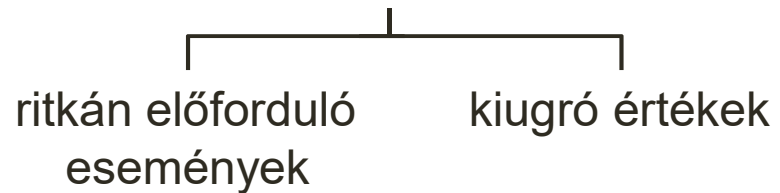
AZ ALFA ALGORITMUS TULAJDONSÁGAI

A bányászathoz csak az **SWF-nets** (Structured Workflow Nets) használható és ennek a hálónak vannak nem implicit helyei.

A következő 2 szerkezet nem használható:



Az Alfa algoritmus nem tud mit kezdeni a **zajokkal**.



MILYEN A JÓ MODELL?

Ahhoz, hogy felfedezzük a jó modellt, feltételezzük, hogy az eseménynapló egy reprezentatív mintája a viselkedésnek.

Azonban két probléma jelenség figyelhető meg:

- **Zaj**: egy esemény log a folyamat tipikus viselkedéseiről nem reprezentatív mintákat szolgáltat (ritka viselkedési minta)
- **Hiányosság**: az eseménynapló nem tartalmaz elegendő eseményt a folyamati szerkezet pontos feltárásához

Megoldás lehet:

- **Implicit helyek** (helyek nem redundások): ártalmatlan és megoldható előfeldolgozással
- **Hurok**: megoldható több módon is (algoritmus változtatása vagy elő/utófeldolgozás)
- **Nem helyi függőség**: alapvető probléma, alfa algoritmus nem használható
- Reprezentáció elfogultsága (nem tudjuk felfedezni az ismétlődő átmeneteket vagy láthatatlan címkéket) egyéb algoritmusok esetén
- Vannak algoritmusok, amelyek a zajra nem érzékenyek
- A hiányosság alapvető probléma, nem tudunk vele mit kezdeni!

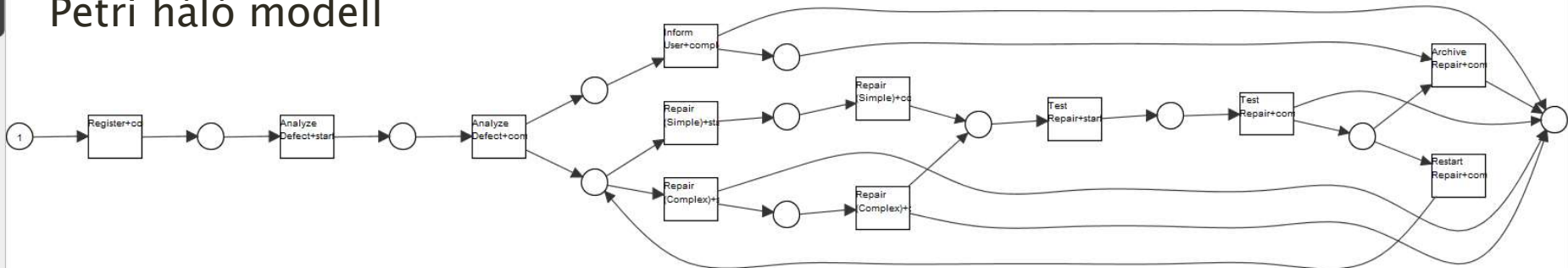
PROM ALPHA ALGORITMUS HASZNÁLATA

The screenshot shows the ProM software interface. At the top, there is a 'ProM' logo and a 'fluxicon' logo. Below the logo, there are navigation icons (back, play, eye) and an 'Activity...' dropdown menu. The main interface is divided into three panels: 'Input', 'Actions', and 'Output'. The 'Input' panel shows a 'DEFAULT XLog' file. The 'Actions' panel is the central focus, displaying a list of actions with a search bar and various icons. The 'Alpha Miner' action is highlighted with a red circle. The 'Output' panel shows 'Petri net' and 'Marking' options. At the bottom of the 'Actions' panel, there are 'Reset' and 'Start' buttons.

Actions List:

- Align Log And Model for Repair (global costs) - D. Fahland (d.fahland@tue.nl) - ModelRepair
- Align Log to Model - D. Fahland (d.fahland@tue.nl) - oma
- Alpha Miner** - S.J. van Zelst, B.F. van Dongen, L.M.A. T. Maer (s.j.v.zelst@tue.nl)
- Animate Event Log in Fuzzy Instance - H. Verbeek (h.m.w.verbeek@tue.nl) - Fuzzy
- Anonymize Event Log - F. Mannhardt (f.mannhardt@tue.nl)
- Anonymize: Obfuscate event attribute names - F. Mannhardt (f.mannhardt@tue.nl)
- Anonymize: Obfuscate literal event attribute values - F. Mannhardt (f.mannhardt@tue.nl)
- Anonymize: Obfuscate literal trace attribute values - F. Mannhardt (f.mannhardt@tue.nl)

Petri háló modell



TOVÁBBI MODELL TÍPUSOK ELŐÁLLÍTÁSA: Mine for a Fuzzy Model

The screenshot displays a software interface for configuring actions. The main window is titled "Actions" and features a central list of actions, each with a play button icon and a search bar. The actions listed are:

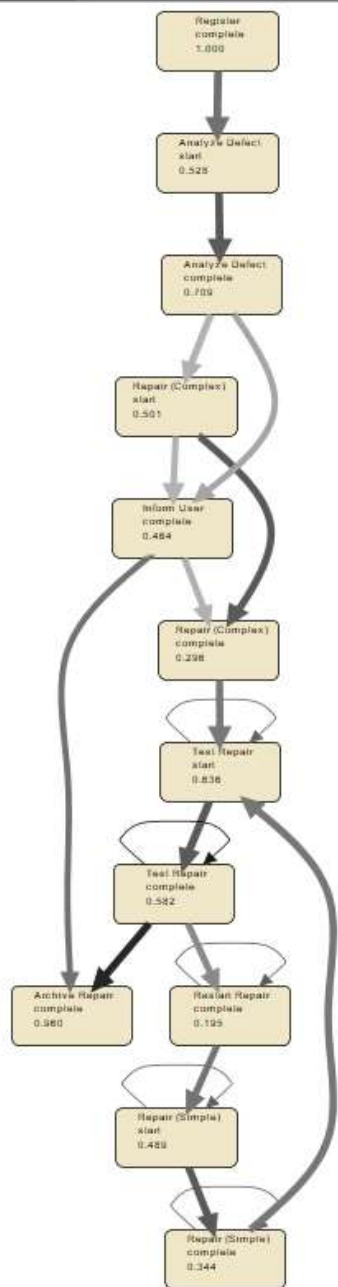
- ILP-Based Process Discovery (S.J. van Zelst)
- ILP-Based Process Discovery (Express) (S.J. van Zelst)
- Interactive Data-aware Heuristic Miner (iDHM) (F. Mannhardt)
- L2Me: Log to Model Explorer (GUI) (x.lu)
- Leemans Episode miner - config dialog (M. Leemans)
- Mine for a Fuzzy Model** (H.M.W. Verbeek) - highlighted in yellow
- Mine for a Heuristics Net using Heuristics Miner (A.J.M.M. Weijters)
- Mine for a Heuristics Net using Heuristics Miner

The interface also includes an "Input" section on the left with a "Click to add input object" button and an "Output" section on the right with a "Click to add output object" button. At the bottom of the central panel, there are "Reset" and "Start" buttons.

Fuzzy Model

PIP

Zoom



Fuzzy Model

PIP

Zoom

