

Process Mining

Descobrint models formals a partir de logs d'un sistema

Jorge Muñoz Gama

Universitat Politècnica de Catalunya

Juny 2009

Contingut

- 1 Motivació
- 2 Process Mining
- 3 Conceptes
- 4 GenetGUI
- 5 Demostració
- 6 Conclusions

Motivació

Desenvolupar una eina, **GenetGUI**, que ajudi en certs aspectes del *Process Mining*.

Process Mining i Procés

Process Mining: tècniques que ens permeten obtenir models formals a partir de la informació dels **processos** enregistrada en els logs d'un sistema.

Process Mining i Procés

Process Mining: tècniques que ens permeten obtenir models formals a partir de la informació dels processos enregistrada en els logs d'un sistema.

Procés: conjunt de passos o accions executades en un determinat ordre per aconseguir una fita.

Procés “Anar al gimnàs”

A l'arribar al gimnàs el primer que es fa són els estiraments. Un cop estirats, el següent pas és fer una mica de bicicleta. Després, si és diumenge, es fa classe de ioga. En canvi, si és un altre dia de la setmana, es fa piscina. I l'últim pas, tan si s'ha fet piscina o ioga, és dutxar-se abans de marxar.

Procés “Anar al gimnàs”

*A l'arribar al gimnàs el primer que es fa són els **estiraments**. Un cop estirats, el següent pas és fer una mica de **bicicleta**. Després, si és diumenge, es fa classe de **ioga**. En canvi, si és un altre dia de la setmana, es fa **piscina**. I l'últim pas, tan si s'ha fet piscina o ioga, és **dutxar-se** abans de marxar.*

Traces i Log d'Events d"Anar al gimnàs"

Dimarts

estirar
bicicleta
piscina
dutxa

Dijous

estirar
bicicleta
piscina
dutxa

Diumenge

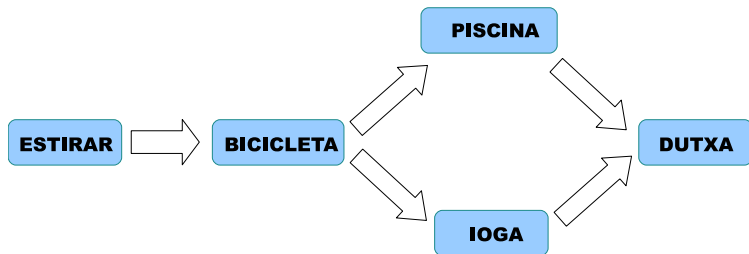
estirar
bicicleta
ioga
dutxa

Model del procés “Anar al gimnàs”

Objectiu del Process Mining: obtenir una descripció del procés a partir dels logs d'events del sistema.

Model del procés “Anar al gimnàs”

Objectiu del Process Mining: obtenir una descripció del procés a partir dels logs d'events del sistema.



Perspectives d'un procés

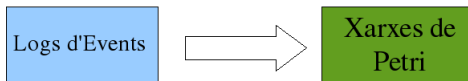
Alguns dels aspectes (o *perspectives*) interessants a obtenir d'un procés:

- **Control de flux:** l'ordre dels events.
- **Xarxa social:** com interactuen les diferents persones. Pot servir per definir *rols* i *grups*.
- **Característiques del funcionament:** informació sobre el funcionament d'un procés.
- **Auditoria i seguretat:** comprovar que no es violi cap restricció de seguretat.
- **Regles de decisió:** descobrir les raons d'una decisió.

Objectiu

L'objectiu és desenvolupar una eina que permeti obtenir un model formal del flux d'un procés a partir dels *Logs d'Events* d'un sistema.

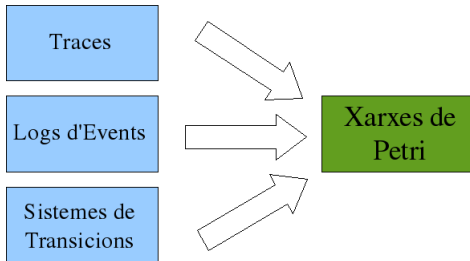
El model formal triat són les *Xarxes de Petri*.



Objectiu

L'objectiu és desenvolupar una eina que permeti obtenir un model formal del flux d'un procés a partir dels *Logs d'Events* d'un sistema.

El model formal triat són les *Xarxes de Petri*.



Traça

- **Traça**: seqüència d'events. L'ordre ve determinat per l'ordre en que han succeït els events.

Traça

- **Traça**: seqüència d'events. L'ordre ve determinat per l'ordre en que han succeït els events.

Exemple de Traça

- estirar - bicicleta - piscina - dutxa

Traces i Logs d'Events

Conceptualment, Traces i Logs d'Events són equivalents.

Traces/Log d'Events: conjunt de traces.

Traces i Logs d'Events

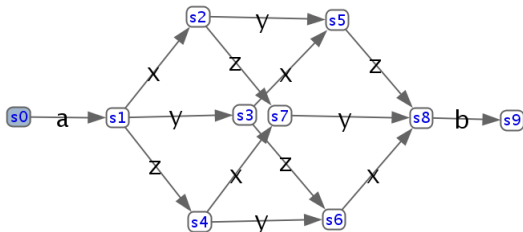
Conceptualment, Traces i Logs d'Events són equivalents.

Traces/Log d'Events: conjunt de traces.

Exemple de Traces/Log d'Events

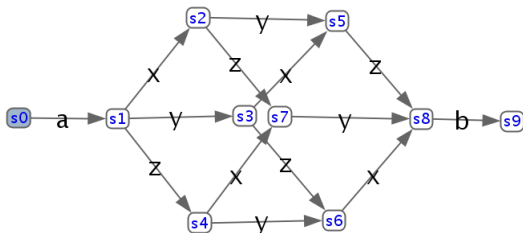
- estirar - bicicleta - piscina - dutxa
- estirar - bicicleta - piscina - dutxa
- estirar - bicicleta - ioga - dutxa

Sistema de Transicions



- Estats
- Accions
- Transicions
- Estat inicial

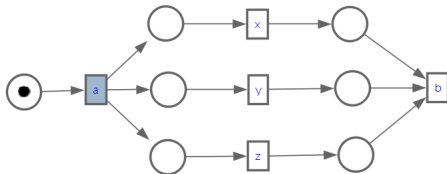
Sistema de Transicions



- Estats
- Accions
- Transicions
- Estat inicial

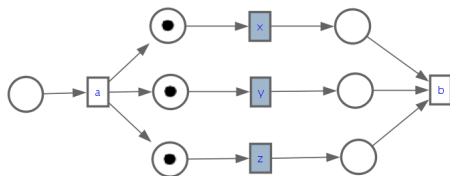
Problema de l'**explosió d'estats** n accions $\rightarrow O(2^n \text{ estats})$

Xarxa de Petri



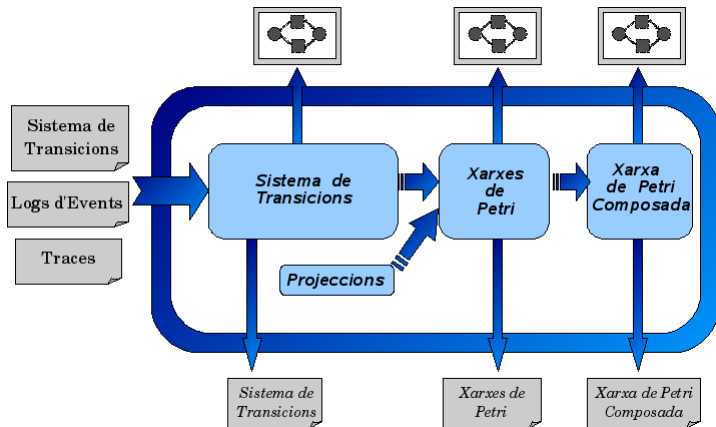
- Places
- Transitions
- Relacions de flux
- Marcatge

Xarxa de Petri

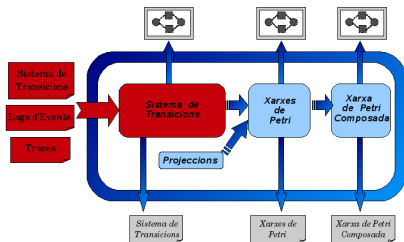


- Places
- Transitions
- Relacions de flux
- Marcatge

GenetGUI



Obrir models



● Formats:

- Log d'Events (MXML)
- Traces
- Sistemes de Transicions

● Llibreries usades:

- Antlr
- Xerxes

Conversió dels models

**LOG
D'EVENTS**

TRACES

**SISTEMA DE
TRANSICIONS**

Conversió dels models



Conversió Log d'Events → Traces

Conceptualment són el mateix (només canvia el format).

Log d'Events (MXML)

```
<ProcessInstance id="0" description="">  
<WorkflowModelElement>x</WorkflowModelElement>  
<WorkflowModelElement>v</WorkflowModelElement>  
</ProcessInstance>
```

Traces

```
x v  
b c j k  
v w x y z  
z x b k
```

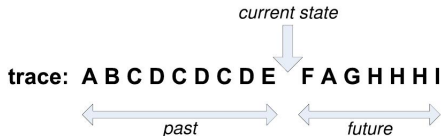
- Una traça per cada *ProcessInstance*
- Un event per cada *WorkflowModelElement*

MXML creat per *Technische Universiteit Eindhoven*

Conversió Traces → Sistema de Transicions

Cal definir els estats:

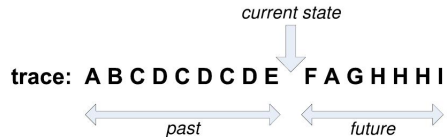
- **Passat / Futur**



Conversió Traces → Sistema de Transicions

Cal definir els estats:

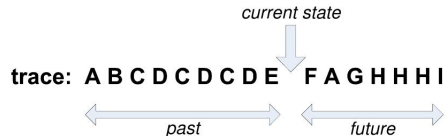
- Passat / Futur
- **Hortizó Màxim**



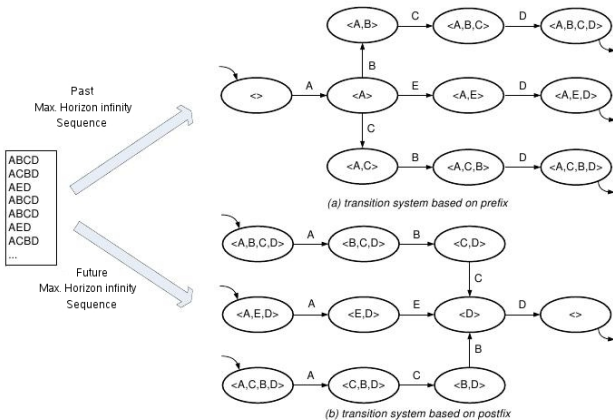
Conversió Traces → Sistema de Transicions

Cal definir els estats:

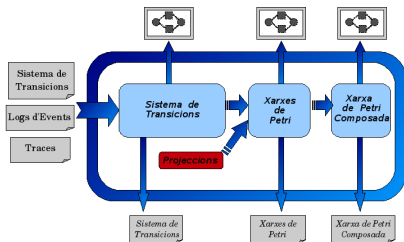
- Passat / Futur
- Hortizó Màxim
- Seqüència / Conjunt



Exemple de conversió



Projeccions



- Particions del Sistema de Transicions
- Cada projecció correspon a una Xarxa de Petri

Algoritmes de Pre-Projecció

Projeccions heurístiques i automàtiques.

GenetGUI proveeix 3 Algoritmes de Pre-Projecció:

- Components Connexos
- Clique
- Clustering

GenetGUI permet afegir nous algoritmes de Pre-Projecció.

Els 3 algoritmes que ofereix GenetGUI es basen en el *Graf de Causalitats*.

Graf de Causalitats

Només es pot generar a partir de *Traces*.

Relacions d'ordenació en Logs

Sigui E el conjunt d'events del log L , i $a, b \in E$ llavors:

- $a >_L b$ si i només si hi ha una traça $\sigma = t_1 t_2 t_3 \dots t_{n-1} i i \in \{1, \dots, n-2\}$ tal que $\sigma \in L$ i $t_i = a$ i $t_{i+1} = b$ (Relació Directa)
- $a \rightarrow_L b$ si i només si $a >_L b$ i $b \not>_L a$ (Relació de Causalitat)

Graf de Causalitats

Graf no dirigit, on els vertexs són els events, i existeix una aresta entre dos events si hi ha una *Relació de Causalitat* entre aquests dos events, sense importar el sentit d'aquesta relació.

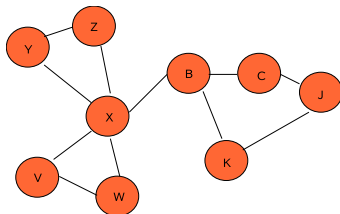
Algoritme de Components Connexos

Idea: grups d'events que no estan relacionats amb altres grups poden ser analitzats de forma separada.

Partir el Graf de Causalitat en components connexos i crear una projecció per a cada un.

Alg. de Components Connexos

```
for all v: Vertices do  
  if not isInComponent(v) then  
    comp := newComponent();  
    vertexSet :=  
    BreadthFirstSearch(v)  
    addVertices(comp, vertexSet)  
  end if  
end for
```



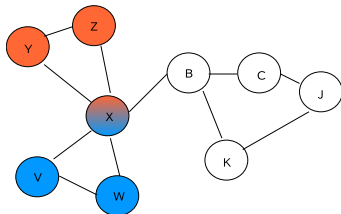
Algorime de Clique

Idea: juntar grups d'events fortament connectats.

Buscar **Cliques** en el Graf de Causalitats.

Clique

Un clique en un graf no dirigit G és un conjunt de vèrtex V tal que per a tot parell de vèrtexs de V , existeix una aresta que els connecta.



Basat en l'*Algorime de detecció de cliques de Bron-Kerbosch*

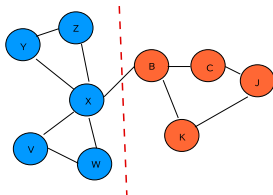
Algoritme de Clustering

Idea: Agrupar conjunts d'events més o menys pròxims, anomenats **Clusters**.

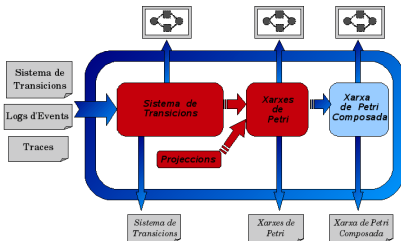
Es basa en el concepte de **Betweenness** d'una aresta: la freqüència en la que aquesta aresta apareix en el camí més curt entre tots els parelles de nodes.

Alg. de Components Connexos

```
for  $1 \leq k \leq numEdgesToRemove$  do  
  computeBetweenness()  
  edge := getHighestBetweenness()  
  removeEdge(edge)  
end for
```

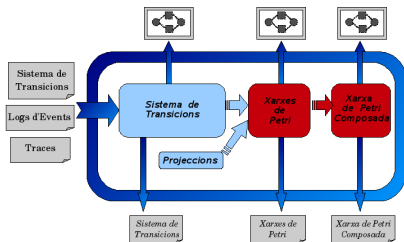


Generar Xarxes de Petri



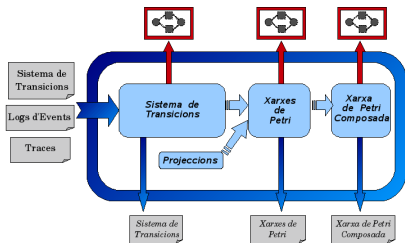
- Generar Xarxes de Petri
- Una xarxa per cada projecció
- Eines usades:
 - Genet [Carmona et al.]

Composar Xarxes de Petri



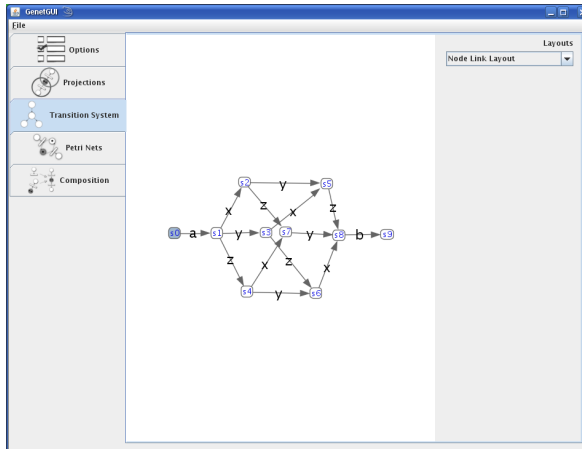
- Crea una única Xarxa de Petri
- Eines usades:
 - PComp [Khomenko]

Visualitzar Models

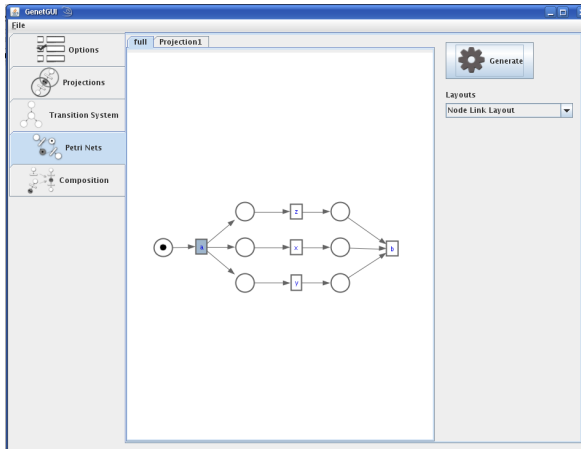


- Visualitzar i interactuar
- Models:
 - Sistema de Transicions
 - Xarxes de Petri
 - Xarxa de Petri composta
- Llibreries usades:
 - Prefuse

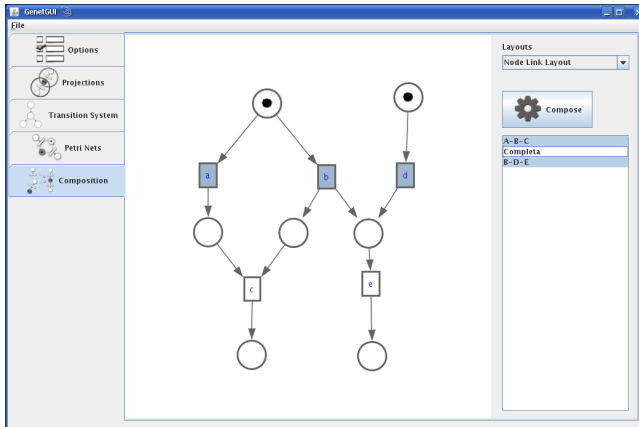
Exemples de Visualitzacions



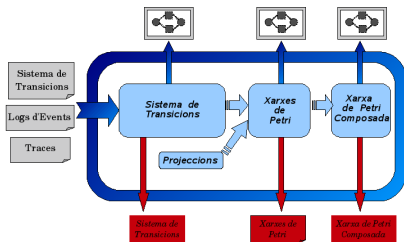
Exemples de Visualitzacions



Exemples de Visualitzacions



Visualitzar Models



- Exportar en mode text
- Models:
 - Sistema de Transicions
 - Xarxes de Petri
 - Xarxa de Petri composta

Demostració

Demostració de GenetGUI

Vídeo de GenetGUI

Conclusions

Objectius assolits:

- Eina gràfica per a *Process Mining*
- Diferents models d'entrada
- Diferents algoritmes de construcció de Sistemes de Transicions
- Algoritmes de Pre-Projecció d'un Sistema de Transicions
- Visualització estàtica i dinàmica de Sistemes de Transicions i Xarxes de Petri

Treball futur:

- Nous models d'entrada, nous layouts i nous algoritmes de Pre-Projecció
- Possible integració en ProM

Moltes gràcies per la seva atenció