Caso di studio in Artificial Intelligence

1 Studenti

Nome Cognome, matricola, mail Federico Canistro, 775605, f.canistro@studenti.uniba.it Ivan De Cosmis, 787066, i.decosmis@studenti.uniba.it

2 Confronto con sperimentazioni precedenti

Nel caso di studio precedente, si é utilizzata la colonna timestamp nel file CSV come identificatore del caso durante la conversione del data frame in un event log. Ció significa che ogni riga nel file CSV veniva trattata come un caso separato con un singolo evento. Di conseguenza, la durata di ogni caso é sempre 0.0 perché é presente solo un evento in ogni caso.

Per calcolare correttamente la durata del caso, é stato necessario specificare una colonna nel file CSV che identifica univocamente ogni caso quando si chiama la funzione $pm4py.format_dataframe$.

Nella nuova versione del codice, stiamo utilizzando la colonna id_exec del file CSV come identificatore del caso quando convertiamo il data frame in un event log. Ció significa che le righe nel file CSV che hanno lo stesso valore per la colonna id_exec , verranno raggruppate nella stessa traccia dentro l' event log.

La differenza rispetto all' event log precedente é che ora le tracce nell'event log possono contenere piú di un evento. Ció consente di calcolare statistiche come la durata del caso, che misura il tempo trascorso tra il primo e l'ultimo evento in ogni traccia.

Ogni output degli algoritmi utilizzati viene salvato in un file pnml e anche in formato png per poterne osservare la sua visualizzazione grafica.

2.1 Funzioni secondarie

get_all_duration_case: controlla la durata di tutti
i case.

compare_pnml_files: Confronta il numero di 'place' e 'transition' nei file.

2.2 Euristic Miner

L'Heuristics Miner é un algoritmo di scoperta di processi piú avanzato che puó gestire meglio le caratteristiche comuni degli event log. Di conseguenza, l'Heuristics Miner é stato in grado di produrre un modello di rete di Petri piú accurato rispetto gli altri algoritmi per lo stesso event log.

2.3 Alpha Miner

L'algoritmo Alpha Miner é un algoritmo di scoperta di processi che puó essere utilizzato per scoprire un modello di rete di Petri da un event log. Tuttavia, l'Alpha Miner ha alcune limitazioni e non é stato in grado di gestire correttamente alcune caratteristiche comuni degli event log, come le attivitá in parallelo e le dipendenze a lungo termine tra le attivitá. In questo caso, l'Alpha Miner ha prodotto un modello di rete di Petri che non rappresenta accuratamente il processo sottostante.

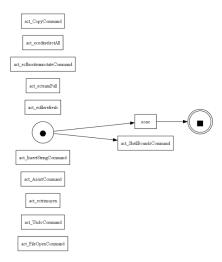


Fig.1. Parte di output alpha sul dataset completo.

2.4 Inductive Miner

L'Inductive Miner é un algoritmo di scoperta di processi che puó essere utilizzato per scoprire un modello di rete di Petri da un event log. In genere, l'Inductive Miner é in grado di gestire event log di grandi dimensioni e complessi in modo efficiente. Tuttavia, in alcuni casi, l'Inductive Miner potrebbe richiedere piú tempo per scoprire un modello di rete di Petri, a seconda delle caratteristiche dell'event log e dei parametri utilizzati. Nel nostro caso l'inductive miner é stato pi1ú di un'ora in esecuzione sul dataset completo e non si é comunque fermato. Mentre testandolo sui dataset separati ha risposto molto bene.

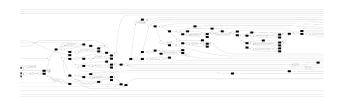


Fig.1. parte di output inductive corretto su uno dei dataset.

Abbiamo inoltre provato ad utilizzare varianti dell'algoritmo Inductive miner, ma sempre senza successo (sempre sul dataset completo). Inoltre, abbiamo anche provato a modificare i parametri dell'algoritmo per migliorarne le prestazioni. Abbiamo utilizzato il parametro noise_threshold per specificare una soglia di rumore per l'algoritmo. L'aumento della soglia di rumore puó aiutare l'algoritmo a gestire meglio gli event log rumorosi e migliorare le prestazioni, ma nel nostro caso ció ha portato ad output scorretti come nel caso dell'alpha miner.

3 Conversione rete di Petri da formato Woman a formato pnml

L'idea di questo algoritmo é leggere da un file pl in input tutti i place, transition e arc nel formato:

- 1. place(X).
- 2. transition(X).
- 3. arc(X,Y).

Vengono salvati all'interno di tre liste indipendenti (conversion_to_pnml.pl, 1):

```
% Dichiaro dinamiche le liste per
% poter essere modificate a runtime
:- dynamic string_list1/1.
:- dynamic string_list2/1.
:- dynamic string_list3/1.
```

In seguito vengono creati i predicati per accedere alle liste (conversion_to_pnml.pl, 10):

```
% Aggiunge stringa nella lista in input (1,2 o 3)
add_string(ListNumber, String):-
   ( retract(string_list(ListNumber, List))
   -> assertz(string_list(ListNumber, [String|List]))
   ; assertz(string_list(ListNumber, [String]))
).

% Restituisce la lista in base al numero in input (1,2 o 3)
get_strings(ListNumber, Strings):-
   ( string_list(ListNumber, Strings)
   -> true
   ; Strings = []
).
```

3.1 Lettura dei dati dal file

3.1.1 process_file(FileName)

La procedura che legge il file e salva i place, le transition e gli arc all'interno delle liste si chiama process_file(FileName).

Questa procedura riceve in input il percorso del file da leggere (conversion_to_petri.pl, 24).

"retractall" viene usato per azzerare tutte le liste ogni volta che viene consultato il file. "process_net" é la procedura che processa la rete.

3.1.2 process_net

In questa procedura troviamo le operazioni che vengono eseguite per estrarre i dati dai fatti letti in input:

- extract_places(NewPlaces): estrae tutti i place dal database e li salva all'interno della lista New-Places
- copy_list(1, NewPlaces): copia la lista in input nella lista dinamica 1
- 3. extract_transitions(NewTransitions): estrae tutti le transition dal database e li salva all'interno della lista NewTransitions
- 4. copy_list(2, NewTransitions): copia la lista in input nella lista dinamica 2
- 5. extract_arcs(NewArcs): estrae tutti gli arc dal database e li salva all'interno della lista NewArcs
- copy_list(3, NewArcs): copia la lista in input nella lista dinamica 3

(conversion_to_pnml.pl, 65)

```
% Estraggo le informazioni dai fatti
process_net :-
    extract_places(NewPlaces),
    copy_list(1, NewPlaces),
    extract_transitions(NewTransitions),
    copy_list(2, NewTransitions),
    extract_arcs(NewArcs),
    copy_list(3, NewArcs).
```

3.1.3 extract_places(Places)

Analizziamo questa procedura che é molto simile a extract_transitions(Transitions) e ex-

tract_arcs(Arcs), cambia solamente la sintassi in base al tipo di dato da estrarre (conversion_to_pnml.pl, 43):

```
% Estraggo tutti i place in una lista extract_places(Places):- findall(Place, place(Place), Places).
```

La procedura findall ricerca tuttii fatti nel formato place(Place) e li salva nella lista Places.

3.1.4 copy_list

Con questa procedura abbiamo due casi (conversion_to_pnml.pl, 55):

- copy_list(IdList, []): Caso in cui la lista in input é vuota e quindi abbiamo finito la copia
- 2. $copy_list(IdList, [H|T])$: Caso in cui la lista bisogna ancora copiarla tutta

```
% Procedura che ci permette di copiare una lista
% in una delle 3 liste dinamiche dato indice in input
copy_list(IdList, []) :-
    true.

copy_list(IdList, [H | T]) :-
    add_string(IdList, H),
    copy_list(IdList, T).
```

Con questa funzione abbiamo finito di analizzare la parte relativa alla lettura dei dati da file in input.

3.2 Scrittura del file pnml

Per scrivere nel file pnml utilizziamo la procedura write_pnml(File, Places, Transitions, Arcs).

3.2.1 write_pnml(File, Places, Transitions, Arcs)

Questa funzione apre il file del percorso inserito in console e inizia a scrivere le varie parti del file pnml (conversion_to_petri.pl, 90).

```
% Scrivo il file pnml dati place, transition e arc
write_pnml(File, Places, Transitions, Arcs):-
open(File, write, StreamWrite),
write_header(StreamWrite),
write_places(StreamWrite, Places),
write_transitions(StreamWrite, Transitions),
write_arcs(StreamWrite, Arcs),
write_footer(StreamWrite),
close(StreamWrite).
```

3.2.2 ask_folder_path(String)

Con questa procedura chiediamo il percoso in cui salvare il file di output (Il percorso alla fine deve includere anche il nome del file e l'estensione pnml) (conversion_to_petri.pl, 38).

```
% Chiedo dove salvare il file pnml
ask-folder-path(String):-
write('Enter path where to save file: ')
read_line_to_string(user_input, String).
```

3.2.3 write_header(StreamWrite)

Questa funzione scrive l'header del file pnml (conversion_to_petri.pl, 101).

```
% Scrivo intestazione file
write_header(StreamWrite):-
write(StreamWrite, '<?xml version="1.0"
encoding="UTF-8"?>'),
nl(StreamWrite),
...
(Abbrevio il codice nella relazione
per problemi di spazio)
```

3.2.4 write_places, write_transitions, write_arcs

Dato che sono tutti molto simili, cambia solamente la formattazione per rispettare il formato pnml, prendiamo in analisi **write_places** Questa procedura ha due casi:

1. Caso base (conversion_to_petri.pl, 112):

```
% Questa procedura serve quando abbiamo % finito tutti i place write.places(StreamWrite, []) :- true.
```

Questo é il caso in cui sono finiti i place da scrivere su file

2. Caso con lista ancora da svuotare (conversion_to_petri.pl, 116):

```
% Questa procedura serve quando
% ci sono ancora place da scrivere
write.places(StreamWrite, [Place|Places]):-
write(StreamWrite, ' <place id="'),
counter(Count),
...
(Abbrevio il codice nella
relazione per problemi di spazio)
...
write.places(StreamWrite, Places).
```

In questo caso ci sono place da scrivere su file e grazie alla funzione write andiamo a scrivere il place rispettando la sintassi del formato pnml. "increment_counter" aumenta il contatore che rappresenta l'id dei place, transition e arc. Infine la funzione richiama se stessa iterando con la lista senza il primo elemento.

3.2.5 write_footer

Questa procedura scrive il pié di pagina del file pnml (conversion_to_petri.pl, 172)

```
% Questa procedura serve quando dobbiamo
% scrivere il pie di pagina pnml
write_footer(StreamWrite):-
    write(StreamWrite, ' </page>'), nl(StreamWrite),
    write(StreamWrite, ' </page), nl(StreamWrite),
    write(StreamWrite, ' </paml>'), nl(StreamWrite).
```

Con questa procedura abbiamo finito di descrivere il codice di questa sezione

3.2.6 Descrizione del caso di test

Questo programma viene caricato in SWI-Prolog usando il comando:
consult('Percorso_fino_al_file/conversion_to_petri.pl').
In seguito viene richiamata la procedura:
process_file('Percorso_fino_al_file/petri_net.pl').
Alla richiesta del percorso dove salvare il file inserire:
Percorso_fino_a_dove_salvare_il_file/nome_file.pnml

Il programma restituirá true e salverá nel percorso indicato il file risutalto. Nel nostro caso nella repository github il file é:

covnersione_petriWoman_pnml/file.pnml

4 Visualizzazioni delle transizioni da un file wf

In questa sezione andiamo a scrivere all'interno di un file di testo le possibili coppie di transizioni successive, in base all'azione di input e alla corrispondenza tra casi. Inoltre vengono anche indicate delle possibili transizioni, che anche se non sono mai state osservate dai casi del sistema, potrebbero comunque avvenire.

4.0.1 Gestione del modulo

Per poter utilizzare le transition caricate abbiamo bisogno di capire il modulo in cui appartengono. Per evitare di fare hardcode con il nome del modulo usato per i test (monday) e quindi forzare sintassi del tipo: monday:transition(Input-Output,Cases,Repetition)

Abbiamo utilizzato una variabile dinamica in cui salviamo l'atomo del modulo letto da tastiera(model_explorator.pl, 1).

```
Soluzione per poter rendere meno vincolante il codice al modulo 7. do_works(Stream): inizia le operazioni di ricerca
% inserito.

% In questo modo basta inserire il nome del modulo da cui sono state

% lette le transition per poter visualizzare nel file i dati

% senza la necessita di fare hardcode col nome del modello
:- dvnamic module / 1.
set_module(Atom) :
      retractall (module(_)),
      assertz (module (Atom)).
get_module(Atom):-
module(Atom).
```

T1 modulo viene chiesto dalla procedura ask_module_name(String) (model_explorator.pl, 127):

```
% Chiedo dove salvare il file pnml
 % Chiedo dove salvale if life plant
ask_module_name(String):-
write('Enter the module name used: '),
read_line_to_string(user_input, String).
```

4.0.2 Inizio delle operazioni

La procedura da cui iniziano tutte le operazioni si chiama **process_file(FileName)**(model_explorator.pl,

15)

```
% Legge il file wf e lo elabora process_file(FileName) :-
          cask_folder_path(Path),
use_module(FileName),
ask_module_name(ModuleString),
string_to_atom(ModuleString, Module),
          set_module (Module),
open (Path, write, Stream),
do_works (Stream).
```

Analizziamo ora le sue operazioni:

- 1. ask_folder_path(Path): chiede da terminale di inserire il percorso in cui andare a salvare il file di output(procedura praticamente identica a quella di lettura del modulo)
- 2. use_module(FileName): carica il modulo del percorso in input
- 3. ask_module_name(ModuleString): chiede da terminale il nome del modulo da usare
- 4. string_to_atom(ModuleString, Module): converte la stringa del modulo in Atomo per poter essere usato durante le operazioni
- 5. set_module(Module): importa la variable dinamica con l'Atomo appena ottenuto
- 6. open(Path, write, Stream): crea il file nel percorso specificato e crea lo Stream da usare

4.0.3 do_works

e analisi delle transizioni

Questa procedura si preoccupa di trovare tutte le transizioni e inserirle in una lista per poterle analizzare in seguito iterando su di essa, sfrutta la procedura **get_module**(model_explorator, 31).

```
% Legge il file wf e lo elabora
List Transitions),
do_work (Module, Stream, List Transitions).
```

Richiama la procedura do_work(Module, Stream, ListTransitions) che inizierá il ciclo per verificare tutte le transizioni.

4.0.4 do_work

Il motivo per cui diamo in input Module é perché ci serve per specificare il modulo nella transizione in input. Questa procedura si occupa di scrivere l'introduzione alla transizione in analisi e poi richiama le procedure per la ricerca delle transizioni da visualizzare. Inoltre itera su se stessa per analizzare tutte le transition in input(model_explorator.pl, 34).

```
% Esegue le operazioni di scrittura su file
e recupero delle informazioni delle transition
do_work(Module, Stream,
[Module:transition(Input-Output,
                                                                  Cases,
Repetition) | Rest]) :-
           write (Stream , '-
           write(Stream,
n1(Stream),
write(Stream, 'Starting transition:'),
n1(Stream),
          nl(Stream),
write(Stream, Input-Output),
write(Stream, ','),
write(Stream, Cases),
write(Stream, ','),
write(Stream, Repetition),nl(Stream),
write(Stream, Repetition),nl(Stream),
writeln(Stream, 'You can continue to the following ones:'),
get_matching_transitions_with_case(Stream,
Output-
                                                                                                Output-_,
                                                                                                Transitions),
          do_work (Module, Stream, Rest).
\% Caso base quando le transizioni da analizzare sono finite do-work (Module, Stream , []).
```

4.0.5 get_matching_transitions_with_case

In questa procedura avviene l'operazione di selezione delle transizioni che seguono a quella in analisi. Viene sfruttato **findall** per poter trovare le transizioni che:

- 1. Hanno l'input cercato
- All'interno della lista dei casi si trova uno dei casi della transizione in analisi

Il secondo punto viene realizzato sfruttando **member** Tutte le transizioni trovate vengono salvate all'interno della lista **NewTransitions**.

Ora peró non abbiamo ancora finito perché abbiamo trovato tutte le transizioni che combaciano con un solo caso della lista(il primo), dobbiamo analizzare gli altri casi. Per fare questo andiamo a reiterare sulla stessa procedura. Per evitare di ottenere duplicati nella lista(transizioni che condividono piú di un caso) usiamo la procedura add_to_list_no_duplicates che leggerá in input la lista delle nuove transizioni, le comparerá con quelle giá trovate(nella prima iterazione non ce ne saranno), e tutte quelle nuove vengono aggiunte alla lista, creando quella nuova che verrá mandata alla prossima iterazione. Tutte le transizioni giá trovate verranno scarate(model_explorator.pl, 49).

Il secondo caso rappresenta la situzione in cui abbiamo finito di analizzare tutte le transizioni riguardante quella corrente, quindi stampiamo il risultato. In questa procedura andiamo anche a richiamare <code>get_matching_transitions</code> che ci permette di ottenere tutte le transizioni che condividono lo stesso input indipendentemente dai casi, sfruttiamo questa procedura per proporre anche transizioni che potrebbero accadere ma che non sono state coperte dai casi del sistema. Per evitare duplicati, dopo aver trovato tutte le transizioni che condividono l'input con quella in analisi andiamo a togliere quelle giá trovate con la procedura <code>add_to_list_deleting_duplicates</code> che date due liste in input restituisce una lista contenente gli elementi della prima senza gli elementi della seconda(eliminando quindi le transizioni valide trovate in precedenza). Infine le stampa.

4.0.6 add_to_list_no_duplicates

Questa procedura riceve in input due lista e ritorna una terza lista che contiene gli elementi della prima lista e della seconda lista senza duplicati(model_explorator.pl, 80).

```
% Questa procedura serve per copiare nella seconda lista tutti
% gli elementi della prima lista che non sono presenti al suo
% interno (Usato per trovare i casi con la corrispondenza senza
% tenere duplicati)
add_to_list_no_duplicates([], List, List).

add_to_list_no_duplicates([H|B], SecondList, Output):-
    member(H, SecondList),
    !,
    add_to_list_no_duplicates(B, SecondList, Output).

add_to_list_no_duplicates([H|B], SecondList, Output):-
    add_to_list_no_duplicates(B, [H|SecondList], Output).
```

Questa procedura ha 3 casi:

- il primo é il caso base in cui la lista di output viene unificata alla seconda
- il secondo caso é dove verifichiamo se l'emento in testa alla prima lista esiste nella seconda, se esiste semplicemente lo saltiamo e reiteriamo
- il terzo caso é il caso in cui l'elemento non apparteneva alla lista, ha fatto fallire il secondo caso e quindi in questo lo andiamo ad aggiungere alla testa della seconda lista e poi reiteriamo.

4.0.7 get_matching_transitions

Questa procedura ci permette di trovare le transizioni che convidono lo stesso input,viene usata per trovare le transizioni che potrebbero accadere ma che non sono state analizzate dal sistema. Usa lo stesso meccanismo della procedura get_matching_transitions_with_case, ma senza il controllo sulla lista dei casi(model_explorator.pl, 73).

4.0.8 add_to_list_deleting_duplicates

Questa procedura é una versione modificata di add_to_list_no_duplicates, perché al posto di andare ad aggiungere ad una lista esistente nuovi elementi, va ad aggiungere ad una lista nuova solo gli elementi che non esistono nella seconda lista(nel nostro caso le transizioni che non convidono casi con quella analizzata, permettendoci di individuare quelle possibili ma non analizzate dal sistema)(model_explorator.pl, 94).

Possiamo trovare 4 casi:

- Il primo caso é quello che riceve la prima lista e la seconda lista in input e richiama la procedura con un nuovo parametro utile a unificare alla fine la nuova lista
- Il secondo caso é il caso base, ovvero quando abbiamo finito gli elementi da confrontare e quindi

la nuova lista viene unificata a quella di output

- 3. Il terzo caso contiene il caso con il controllo che verifica se esiste o no l'elemento nella seconda lista, se esiste va avanti e itera di nuovo, altrimenti passa al quarto caso
- 4. L'ultimo caso aggiunge l'elemento trovato alla nuova lista e reitera

4.0.9 print_list

Con questa procedura stampiamo una lista in input nello Stream ricevuto(model_explorator.pl, 113).

```
% Procedura che stampa la lista in input e divide con dei trattini % la formattazione print_list(Stream, []) :- write(Stream, '- '), nl(Stream), flush-output(Stream). print_list(Stream, [H|B]) :- writeln(Stream, H, , print_list(Stream, B).
```

Possiamo trovare 2 casi:

- Il primo caso é quello che considera la lista vuota e quindi la fine della stampa che viene indicata con dei trattini orizzontali, in seguito viene svuotato anche il buffer altrimenti si rischia di riempirlo e non riuscire piú a stampare su file
- Il secondo caso é il caso con la lista ancora con elementi da stampare, viene stampata la testa della lista e poi viene reiterato il ciclo.

Con questa procedura concludiamo l'analisi del codice di questa sezione

4.0.10 Descrizione del caso di test con monday.wf

Questo programma viene caricato in SWI-Prolog usando il comando:
consult('Percorso_fino_al_file/model_explorator.pl').
In seguito viene richiamata la procedura:
process_file('Percorso_fino_al_file/monday.wf').
Alla richiesta del percorso dove salvare il file inserire:
Percorso_fino_a_dove_salvare_il_file/nome_file.txt

Alla richiesta del nome del modulo inserire il nome del

 ${\it modulo\ caricato:}$

 monday

Il programma restituirá true e salverá nel percorso

indicato il file risutalto. Nel nostro caso nella repository github il file é:

 $visualizzazione_modello/risultato.txt$