UNIVERSITÄT KOBLENZ · LANDAU

Business Process Management

Exercise 7

Group 04

Rahul Bhanushali

Davud Ismayilov

Chanchal Gopalakrishnan

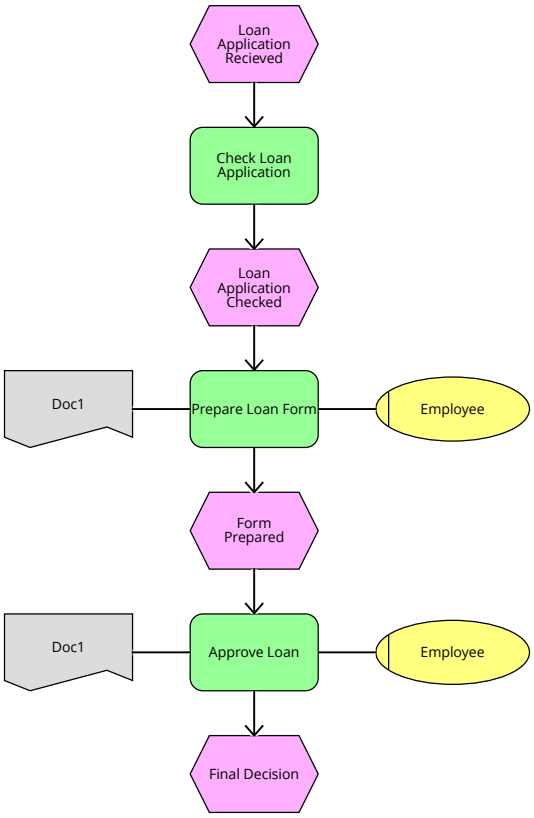
Harita Sawant

Vishal Kumar

Michael Agyekum Bremang

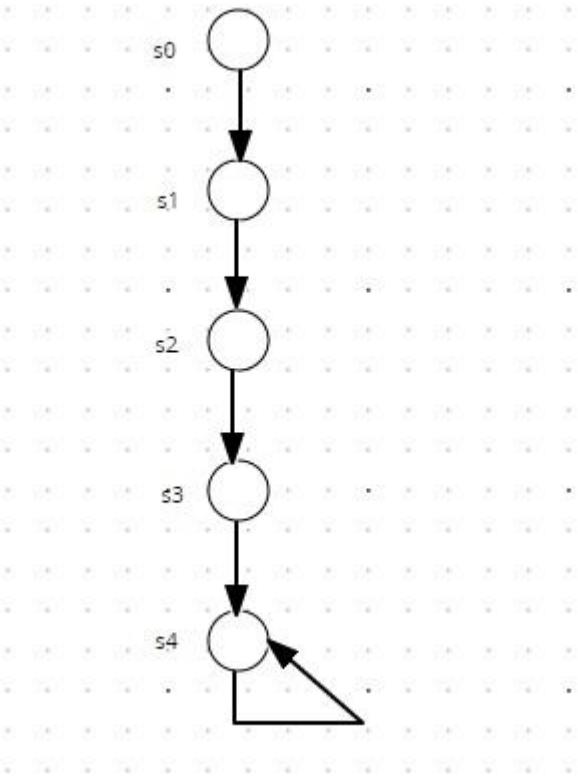
TOTAL: 7,25/10

Ex_7


Loan Application Recieved
Check Loan Application
Loan Application Checked
Doc1
Prepare Loan Form
Employee
Form Prepared
Doc1
Approve Loan
Employee
Final Decision

Exercise 7

Task 1 4,5/7



L(s0) = {start},
L(s1) = {n_check_loan_application},
L(s2) = {n_prepare_loan_form, o_employee, d_doc1},
L(s3) = {n_approve_loan, o_employee, d_doc1},
L(s4) = {end}

Error 1 : Modeling Error : In the EPC diagram, after the loan application is received, it is being checked before sending it to employee for preparing loan form but checking the loan application does not have any effect on the overall process. This type of modeling error is called **Decision Without Impact**.

// To detect this error in CTL, we can form a query that next state of start should be prepare loan and checking loan application will be avoided

What if the check happened after the preparation? Then this would still be a weakness (-0,25)

M, s0 |= AG(start → AX(n_prepare_loan_form))

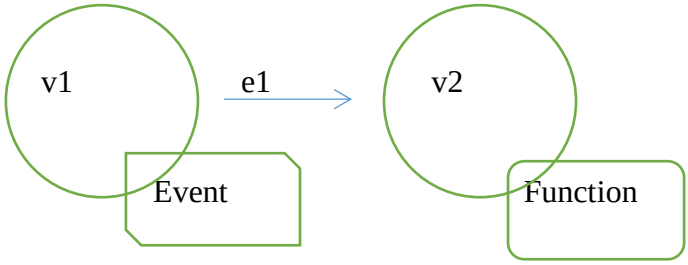
This would always return true, unless every single element in your process is called "start"(-0,5)

To check for all functions directly following start, you could just have written M,s0 |= AX(n_prepare_loan_form)

// To detect this error in GMQL, we can form a pattern that once loan application is received, it must be immediately followed by preparing loan form avoiding checking

AdjacentSuccessors(
ElementsWithAttributeOfValue(V, label, "*Received"),
ElementsWithAttributeOfValue(V, label, "Prepare Loan*")
)

// We can detect this error in DMQL by restricting number of edges, defining captions etc



P(v1) : VID = {v1}, VCAPTION = {"Loan Application Received"},
VTYPES = {Event}
P(v2) : VID = {v2}, VCAPTION = {"Prepare Loan Form"}, VTYPES = {Function}

P(e1) : EID = e1, ECAPTION = " ", MINL = 1, MAXL = 1, DIR = org,
MINVO = MAXVO = MINEO = MAXEO = 0,
VTYPESR = ETYPESF = VTYPESF = ETYPESR = {}

Error 2 : Organization Error : Since the EPC model is model of a loan process so it must follow 4 eye principle. One of the condition of 4 eye principle is that different tasks must be performed by different organization units i.e a different person should approve the loan than the person who prepared the loan form but in our EPC, instead of a manager, employee is approving the loan and also performing all the tasks which raises the error of **Unclear Responsibility** since the same employee has multiple responsibilities.

// We can detect this error in CTL by finding out path where employee approves the loan What if a different person prepared the loan form? Then it would be ok for the employee to check it. (-0,5)

M, s0 $\models \neg AF (n\_approve\_loan \wedge o\_employee)$

//For GMQL, we can detect a path where loan is not approved by employee

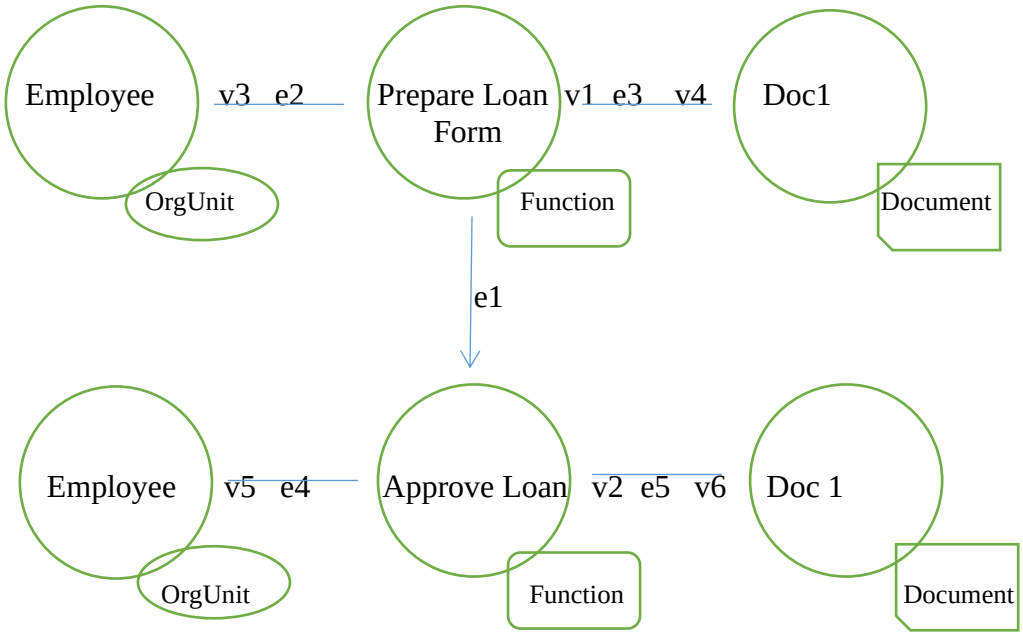
PathsNotContainingElements(
ElementsWithAttributeOfValue(V, label, “Approve Loan”),
ElementsOfType(V, OrgUnit),

ElementsWithAttributeOfValue(V, label, “Employee”)

)

In your case the path is of length 1, so it cannot contain elements anyways (-1)

//For DMQL, we can define a global rule that caption value of both the functions should not be same



P(v1) : VID = {v1}, VCAPTION = {“Prepare Loan Form”}, VTYPES = {Function}
P(v2) : VID = {v2}, VCAPTION = {“Approve Loan”}, VTYPES = {Function}
P(v3) : VID = {v3}, VCAPTION = {“Employee”}, VTYPES = {Orgunit}
P(v4) : VID = {v4}, VCAPTION = {“Doc1”}, VTYPES = {Document}
P(v5) : VID = {v5}, VCAPTION = {“Employee”}, VTYPES = {Orgunit}
P(v6) : VID = {v6}, VCAPTION = {“Doc1”}, VTYPES = {Document}

P(e1) : EID = e1, ECAPTION = “ ”, MINL = 1, MAXL = -1, DIR = org,
MINVO = MAXVO = MINEO = MAXEO = 0, VTYPESR = {Event},
ETYPESF = VTYPESF = ETYPESR = {}

P(e2) : EID = e2, ECAPTION = “ ”, MINL = 1, MAXL = 1, DIR = none,
MINVO = MAXVO = MINEO = MAXEO = 0, VTYPESR = {~~Orgunit~~},
ETYPESF = VTYPESF = ETYPESR = {}

A path of length 1 cannot contain any elements (-0,25)

P(e3) : EID = e3, ECAPTION = " ", MINL = 1, MAXL = 1, DIR = none,
MINVO = MAXVO = MINEO = MAXEO = 0, VTYPESR = {~~Document~~},
ETYPESF = VTYPESF = ETYPESR = {}

P(e4) : EID = e4, ECAPTION = " ", MINL = 1, MAXL = 1, DIR = none,
MINVO = MAXVO = MINEO = MAXEO = 0, VTYPESR = {~~Orgunit~~},
ETYPESF = VTYPESF = ETYPESR = {}

P(e5) : EID = e5, ECAPTION = " ", MINL = 1, MAXL = 1, DIR = none,
MINVO = MAXVO = MINEO = MAXEO = 0, VTYPESR = {~~Document~~},
ETYPESF = VTYPESF = ETYPESR = {}

Global Rules :

// Organization units of both functions should be different :

[v3].[Caption].value != [v5].[Caption].value

Task 2 2,75/3

1. Exploring Interpretability for Predictive Process Analytics

This is not the title of the paper

Summary-

Modern predictive analysis make use of different statistical information,data and machine learning technique for decision making. Historical data is used to make prediction for future and make decision based on it.In business process predictive analysis plays an important part, as all stored historical data would be taken into consideration if prediction about future scenarios of business needs to be met.The main purpose is providing the best assessment of what will happen in future based on what has happened.There are different models used to make prediction and this model needs to be trained to work on the data and give the output.In this paper we are focused on machine learning models and why interpretability is important to incorporate to the models.Interpretability is something by which humans would be able to understand why particular prediction is made.

What are the results of the paper? (-0,25)

the research question of the paper

There are different machine learning models or techniques which needs to be trained with various input features. Generally we expect the output or prediction to be accurate but thats not sufficient. The model needs to be interpretable for humans to understandable.Paper generally tells about necessity of intepretability in machine learning techniques.

the research paradigm/methodology of the work:

Predictive process monitoring approach was used here,where we have sequence of data and from which all the data are grouped together based on their characteristics and properties together known as bucketing .Different bucketing, encoding and supervised learning algorithm was applied to the dataset. The data are then encoded.This features grouped together are then used to train predictive model.To incorporate interpretability,post hoc interpretation was applied to it.
Different analaysis was found based on the research on real life example:

- Predictive techniques using single bucket may present higher accuracy than prefix length bucket.
- In encoding the data, aggregate encoding would be the better option when we are working on business process.

the evaluation methods of the work:

Different evaluation methods was derived from the research.
-Model interpretation is quite useful when choosing predictive model.
-When dealing with business process, its quite necessary to have domain knowledge for better understanding of model.
-Interpretation can help better understanding of predictive model.
-It would be helpful for people in business if particular predictive model could clarify why the prediction was made.

whether you see any limitations that are not mentioned in the paper

As data is skewed , its hard dealing with it. We might have duplicate data , or missing values. Its time consuming to deal with such data and make it ready to implement predictive modeling.
When we are unable to find creative solution to deal with data.
Sometimes wrong evaluation metrics are used to deal with data.

2. Extracting Annotations from Textual Descriptions of Processes

Summary:
Business Process Management requires process models to have system aided solution to variety of tasks. But it requires efforts and knowledge to understand the modelling languages, this is why the companies prefer to have informal textual process descriptions. The authors in this study have proposed a novel technique to extract ATDP (Annotated textual descriptions of processes) specifications from textual descriptions following a formal scheme and being human-readable. They have discussed the related work and the main focus of comparison of their work which is the state-of-the-art technique. They have also discussed the preliminaries of the paper which is the Freeling NLP software and its language processing layers, the ADTP process and the Tregex patterns used in the work.

Basically, the approach in this paper follows the cycle where the NLP analysis is first performed on the textual process description received and the dependency tree for each sentence is extracted. Later this tree is transformed in to format suitable for Tregex patterns and identifies potential activity fragments by labelling them as <ACTION>. In the next step, process entities fragments are extracted followed by the extraction of activity fragments. Lastly the relation between the activity nodes is extracted by the final set of patterns and we have our ADTP.
To evaluate the approach there are two experiments done namely activity extraction and relation extraction and the performance is compared by the gold standard corpora which is manually annotated by expert curators to test the automated text processing systems. At the end of the experiment, we see that the approach is low in terms of recall value but it is high in terms of precision and better in terms of F1 score compared to the other work. Unlike the existing approaches they have used the Tregex query language patterns that help to find the substructures within the syntactic dependency trees corresponding to the NLP analysis of each sentence. This approach has resulted in better accuracy in detection of main process elements.

3. Looking for Meaning: Discovering Action-Response-Effect Patterns in Business Processes

Jelmer J. Koorn[1], Xixi Lu[1], Henrik Leopold[2,3], and Hajo A. Reijers[1]

Authors of this paper try to overcome the challenges in Process mining techniques and to represent the insights of the mining results in understandable manner.
This research paper provides the technique called 'Action-Response-Effect' to improve the Process mining techniques by providing the useful insights from the mining results. Main aim of this technique is to determine the possible outcomes as Effect corresponding to Action and Response. This technique aims to solve two challenges that are to 'represent each action-response connectivity in such a way to determine the effect corresponding to such relations' and 'to filter certain responses which allows to gain useful information from the process model.
Formalization techniques are used to implement the algorithm to discover logs of action-response events. Then such logs are then processed to fetch effects of the responses. Here span between response and its corresponding effect impacts the reliance between them. To check the dependency between effect and response Chi-Square statistics is used. If Chi-square value is more, the more likely that effect is dependent on response otherwise it is not dependent.
After finding dependency for each relation, it is displayed in graphical representation. Number of graphical representations is determined by total number of actions where Chi-square value is significant.
Health-care data set was used to perform such action-effect-response technique to find the patterns to determine effect of each treatment to help in decision making between multiple treatments for a disease.