

SoftCOM 2016

Diana Haliță

September 22, 2016

Babeș-Bolyai University, Cluj-Napoca, România

Formal Concept Analysis

Is Formal Concept Analysis suitable to improve Electronic Health Record Systems?

Analysing the Effect of Changing the Educational Methods by Using Formal Concept Analysis

Formal Concept Analysis

Is Formal Concept Analysis suitable to improve Electronic Health Record Systems?

Analysing the Effect of Changing the Educational Methods by Using Formal Concept Analysis

Formal Concept Analysis

Is Formal Concept Analysis suitable to improve Electronic Health Record Systems?

Analysing the Effect of Changing the Educational Methods by Using Formal Concept Analysis

FORMAL CONCEPT ANALYSIS

SoftCOM 2016

Diana Haliță

September 22, 2016

Babeș-Bolyai University, Cluj-Napoca, România

PREREQUISITES

Definition

Formal Context: triple $\mathbb{K} = (G, M, I)$ with

- set G of objects
- set M of attributes
- incidence relation $I \subseteq G \times M$

CROSS-TABLE REPRESENTATION

Planets	Near	Far	Small	Medium	Large	s-yes	s-no
Mercury	x		x				x
Venus	x		x				x
Earth	x		x			x	
Mars	x		x			x	
Jupiter		x			x	x	
Saturn		x			x	x	
Uranus		x		x		x	
Neptun		x		x		x	
Pluto		x	x			x	

Definition

Formal Concept: pair (A, B) with

- extent $A \subseteq G$
- intent $B \subseteq M$
- $A \times B \subseteq I$
- maximal w.r.t. this property

The planet context has the following concepts:

- (Φ ; {Near; Far ; Small ; Medium; Large; s-yes; s-no})
- ({Jupiter; Saturn}; {Large; Far ; s-yes})
- ({Uranus; Neptun}; {Medium; Far ; s-yes})
- ({Earth; Mars}; {s-yes; Near; Small})
- ({Mercury; Venus}; {s-no; Near; Small})
- ({Jupiter; Saturn; Uranus; Neptun}; {Far ; s-yes})
- ({Earth; Mars; Mercury; Venus}; {Near; Small})
- ({Jupiter; Saturn; Uranus; Neptun; Earth; Mars}; {s-yes})
- ({Mercury; Venus; Earth; Mars; Jupiter; Saturn; Uranus; Neptune}; Φ)

CROSS-TABLE REPRESENTATION

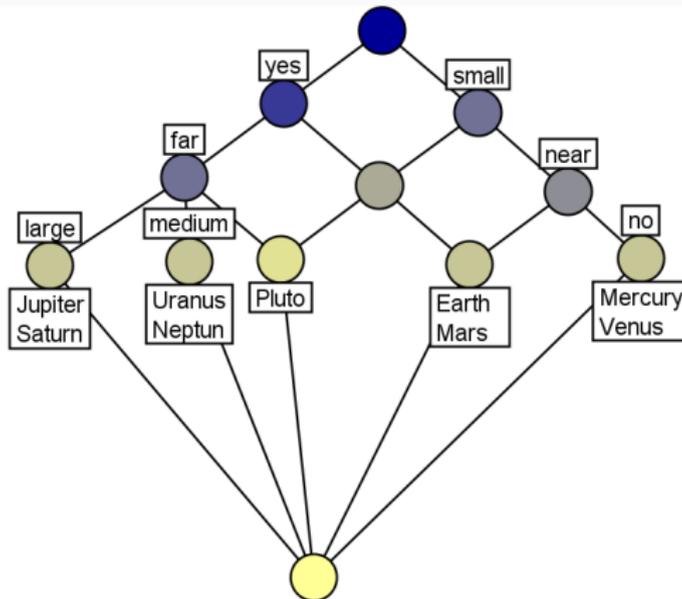


Figure: Concept lattice of the formal context

- Triadic FCA was introduced by F. Lehman and R. Wille in 1995
- Polyadic FCA was introduced by G. Voutsadakis in 2002

Definition

n-context: $(n+1)$ -tuple $\mathbb{K} = (K_1, \dots, K_n, R)$ with

- K_1, \dots, K_n sets
- n-ary incidence relation $R \subseteq K_1 \times \dots \times K_n$

n-concept: n-tuple (A_1, \dots, A_n) with

- $A_1 \times \dots \times A_n \subseteq R$
- maximal w.r.t. this property

TRIADIC CONTEXT EXAMPLE

Triadic setting:

- users as objects
- chains of pages as attributes
- timestamp as conditions

w3	A	B	C	D
LT-LA	X			
LT-LE	X	X	X	

w4	A	B	C	D
LT-LA	X			X
LT-LE	X	X		

Triconcepts:

- $(\{LT - LA, LT - LE\}, \{A\}, \{w3, w4\})$,
- $(\{LT - LE\}, \{A, B, C\}, \{w3\})$,
- $(\{LT - LE\}, \{A, B\}, \{w3, w4\})$,
- $(\{LT - LA\}, \{A, D\}, \{w4\})$,
- $(\emptyset, \{A, B, C, D\}, \{w3, w4\})$
- $(\{LT - LA, LT - LE\}, \{A, B, C, D\}, \emptyset)$.

IS FORMAL CONCEPT ANALYSIS SUITABLE TO IMPROVE ELECTRONIC HEALTH RECORD SYSTEMS?

SoftCOM 2016

Diana Haliță, Christian Săcărea

September 22, 2016

Babeș-Bolyai University, Cluj-Napoca, România

Motivation

Prerequisites

Data Collection and Analysis

Conceptual Knowledge Processing at Work

Conclusions and Future Work

Motivation

Prerequisites

Data Collection and Analysis

Conceptual Knowledge Processing at Work

Conclusions and Future Work

Motivation

Prerequisites

Data Collection and Analysis

Conceptual Knowledge Processing at Work

Conclusions and Future Work

Motivation

Prerequisites

Data Collection and Analysis

Conceptual Knowledge Processing at Work

Conclusions and Future Work

Motivation

Prerequisites

Data Collection and Analysis

Conceptual Knowledge Processing at Work

Conclusions and Future Work

MOTIVATION

- FCA's efficient algorithms and expressive power
⇒ insight on knowledge structures;
- **Purpose:** improving EHR systems with new, FCA grounded features;
- **Strategy FCA:**
 - build 'knowledge landscapes';
 - use them in the framework of EHR;
 - use triadic approach;⇒ improve the task of knowledge discovery in EHR systems;
- **Why FCA?**
 - most of the related work is focused on research.
 - there is less discussion on how state of the art knowledge discovery and representation techniques can improve EHR systems and aid clinicians.

PREREQUISITES

- EHR systems provide alerts, decision support and data analysis
⇒ care management and research;
- EHR = a repository of electronically maintained information about a patient's health status and health care;
- **Why we use FCA?**
 - well known for its effective knowledge discovery algorithms;
 - provide a solid communication basis for further data analysis;
 - extract valuable knowledge from medical data sets;
- **Expected results:**
 - solve the knowledge discovery, processing and representation task in EHR systems.

DATA COLLECTION AND ANALYSIS

- **2 medical datasets:**
 - a cancer registry;
 - an adverse drug reaction database;
- collected at the Ion Chiricuta Oncological Institute of Cluj-Napoca;
- 6000 patients records, with roughly 100 attributes;
- **Information:**
 - about the patient (ID, age, date of birth, date of death);
 - about medical procedures, treatments, results and adverse drug reactions;
- **Tools:**
 - ToscanaJ system is used to build conceptual scales and to visualize knowledge clusters.

CONCEPTUAL KNOWLEDGE PROCESSING AT WORK

- Integrated view of patient data;
- Clinical decision support;
- Clinician order entry;
- Access to knowledge resources;
- Integrated communication and reporting support.

Purpose

- improve the integrated view of patient data component by means of FCA;
- give to the clinician a more comprehensive view of the analyzed data;

- EHR systems provide individual patient views: medical problems, treatment, allergies, reminders;
- there is no comprehensive view of the medical data, allowing the clinician to compare, browse, navigate, and explore the entire dataset in a simple, yet intuitive way;
- **Our proposal:**
FCA enhanced views - landscapes of knowledge
 - patient;
 - disease;
 - treatment.

BROWSING SCENARIOS

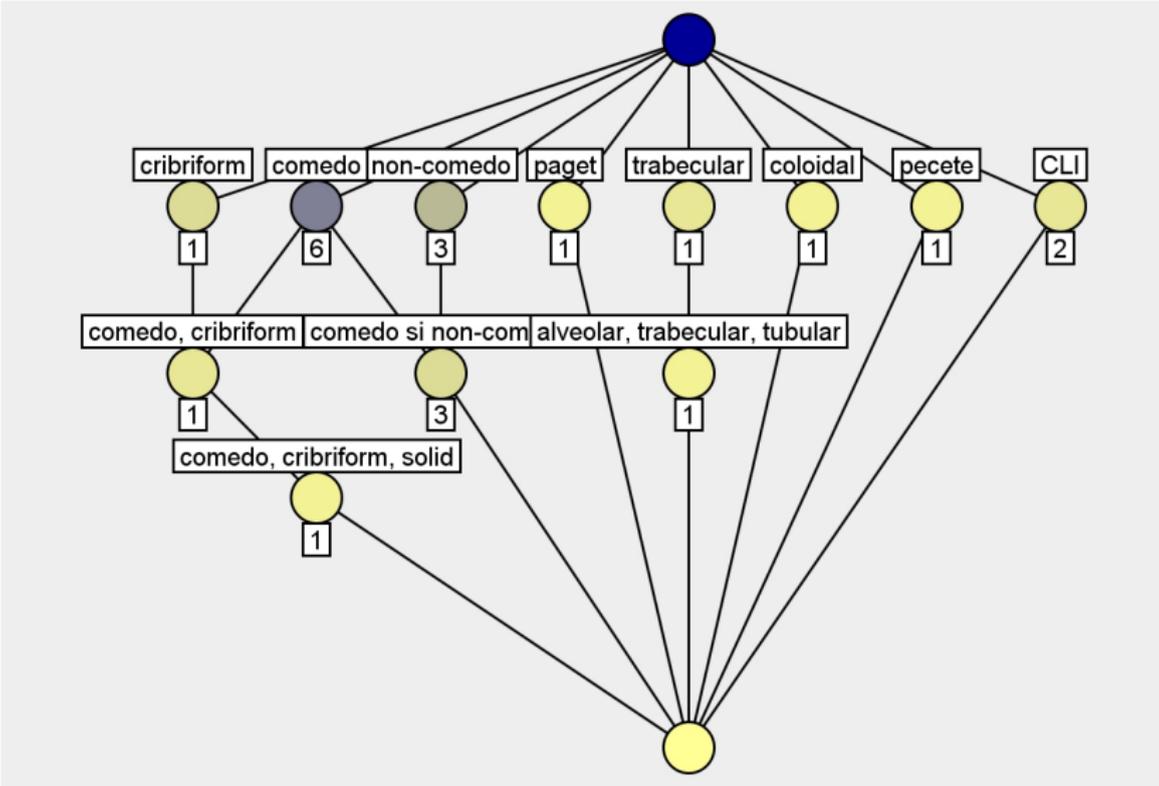


Figure: Carcinoma map

BROWSING SCENARIOS

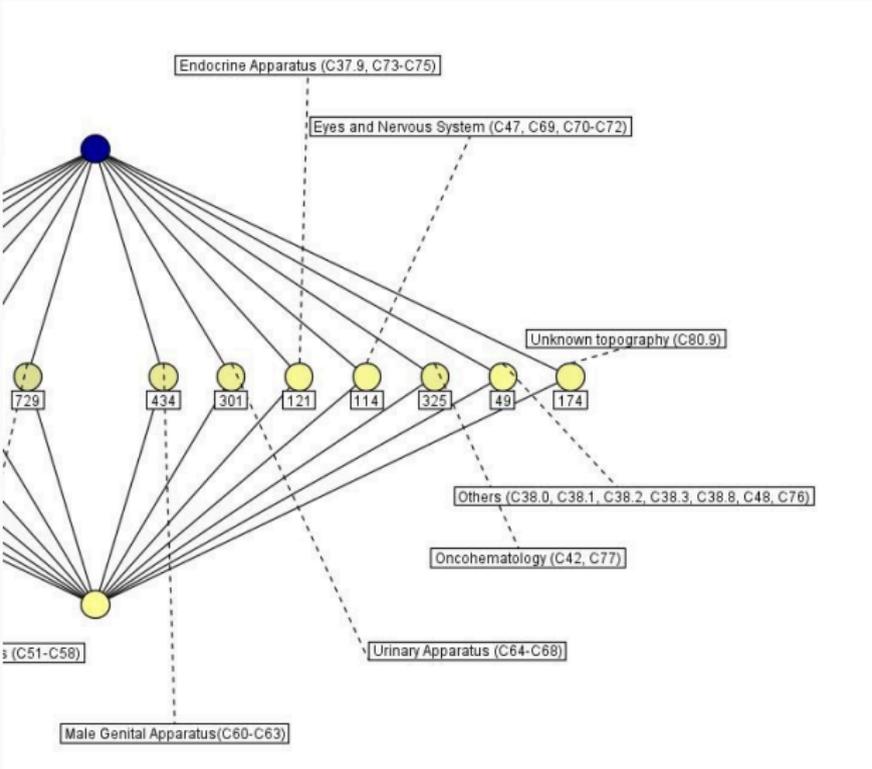


Figure: Cancer tumor topography. Partial view

What is the impact of the age and of the primary tumor location on the intensity and distribution of the adverse effects?

- select the patient age scale;
- select the location of the tumor;
- select the adverse effects;

ANALYZING ADVERSE DRUG REACTIONS

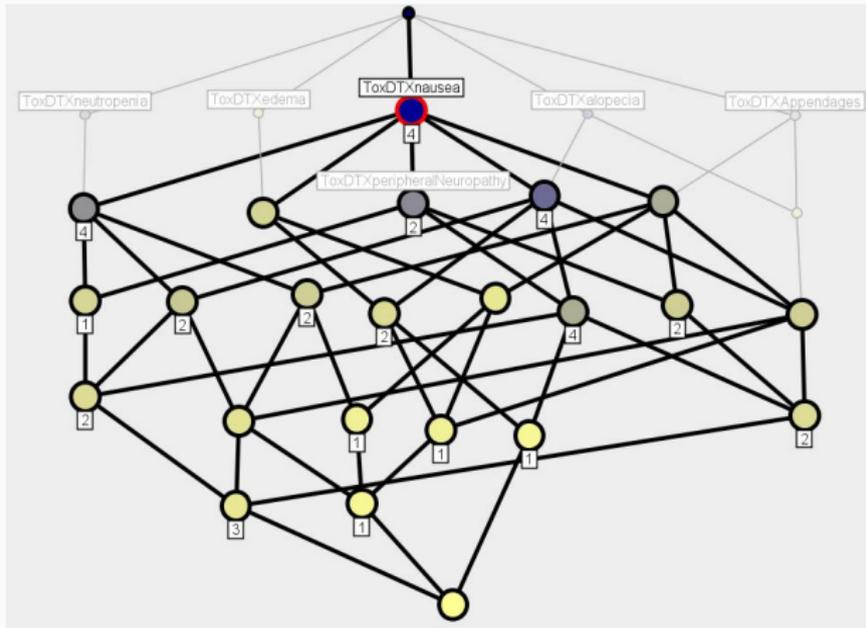


Figure: DTX adverse reactions: complete distribution view. The Nausea node is selected, highlighting all combinations of adverse effects that include nausea. The results are in count format, showing the number of patients that have suffered all the adverse effects determining a certain node.

Which is the perspective on survival rate considering different age categories? Which are the types of cancer exposed?

- select the Age scale;
- select the Survival rate;
- select the Types of cancer.

Which is the status of the patient (alive/dead) considering a specific survival period and a curative or non-curative treatment?

- select the Types of treatments scale;
- select the Survival Rate;
- select the Vitality;
- select the Cause of death.

Which is the strongest type of cancer?

- select the Types of treatments scale;
- select the Survival Rate;
- select the Vitality;
- select the Cause of death;
- select the Types of cancer;
- select the Morphology.

Hormonal treatment for the patients with a tumor on the male genital apparatus

Although this treatment is not considered a curative treatment it has a success rate in vitality for about 78% of the patients.

3FCA applied in medical data has the potential to offer a global view over the data and allows to navigate it in a comprehensive manner.

Tricontext 1

- TYPES OF CARCINOMA (objects)
- FIRST APPLIED TREATMENT (attributes)
- CONFIRMATION MODE (conditions)

[(non-comedo);(PCTneoadj);(Pc T,Pc N)]

[(comedo);(surgery, PCTneoadj);(Bio)]

[(cribriiform,comedo,alveo,non-comedo,paget);(PCTneoadj);(Bio)]

[(comedo,alveo,cli);(surgery);(Bio)]

[(comedo,non-comedo,coloidal,cribriiform); (PCTneoadj);(Pc T)]

[(comedo,cribriiform);(PCTneoadj);(Pc T,Bio)]

Tricontext 2

- CONFIRMATION MODE (objects)
- FIRST APPLIED TREATMENT (attributes)
- COMITEE INDICATION (conditions)

[(Pc T);(PCTneoadj);(cons,rad,nosurgery)]

[(Pc T,Bio);(surgery,PCTneoadj);(cons)]

[(Bio);(surgery,PCTneoadj);(rad)]

[(Bio);(surgery,PCTneoadj);(cons, rad)]

[(Pc T,Bio);(PCTneoadj);(cons, rad)]

[(Pc T);(PCTneoadj,Meta);(rad)]

[(Pc T,Bio,Pc N);(PCTneoadj);(rad)]

Tricontext 3

- LOCATION OF METASTASIS (objects)
- TREATMENT (attributes)
- DRUG ADVERSE REACTIONS (conditions)

[(liver);(AC,DTX,TXT);(vomiting,nausea)]

[(bone);(EC,CMF);(vomiting,alopecia,neutropenia,nausea)]

[(pleurisy);(AC,TXT);(vomiting, alopecia, nausea)]

[(lung);(AC); (edema, appendages, neutropenia, nausea)]

VISUALIZING DATA STRUCTURE IN A CIRCULAR LAYOUT

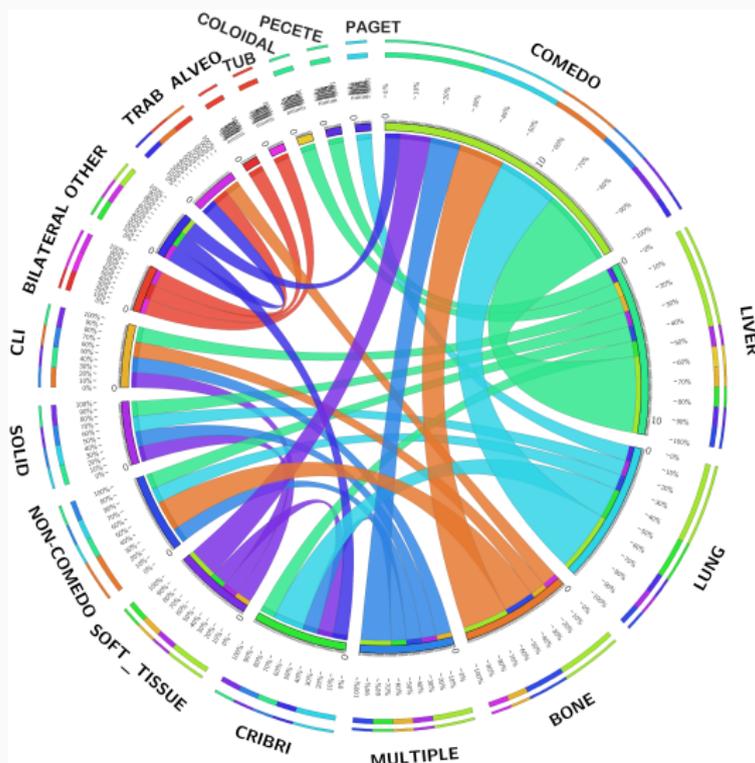


Figure: Carcinoma map vs metastasis location in a circular layout

VISUALIZING DATA STRUCTURE IN A CIRCULAR LAYOUT

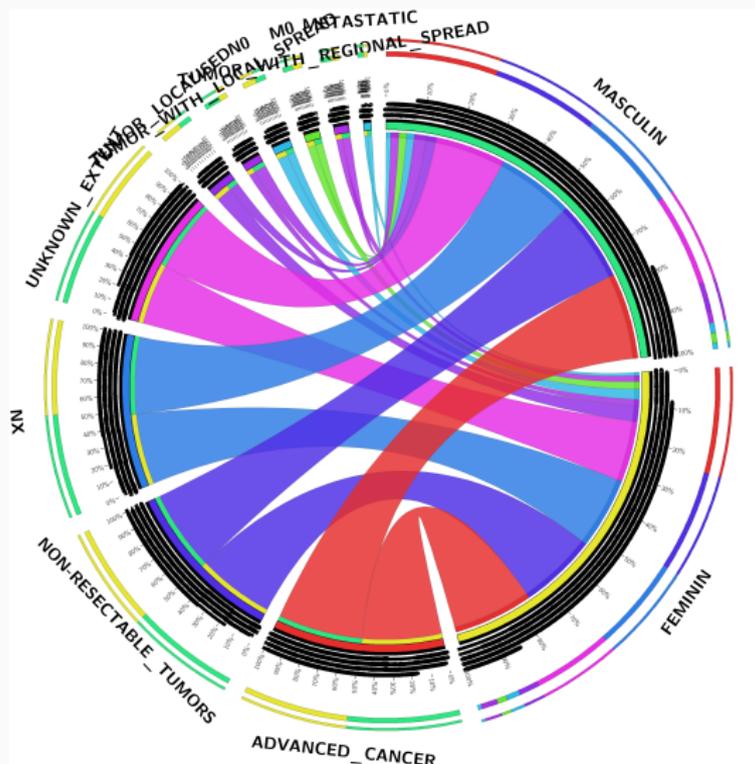


Figure: Tumor node metastasis (TNM) staging system vs patient sex

VISUALIZING DATA STRUCTURE IN A CIRCULAR LAYOUT

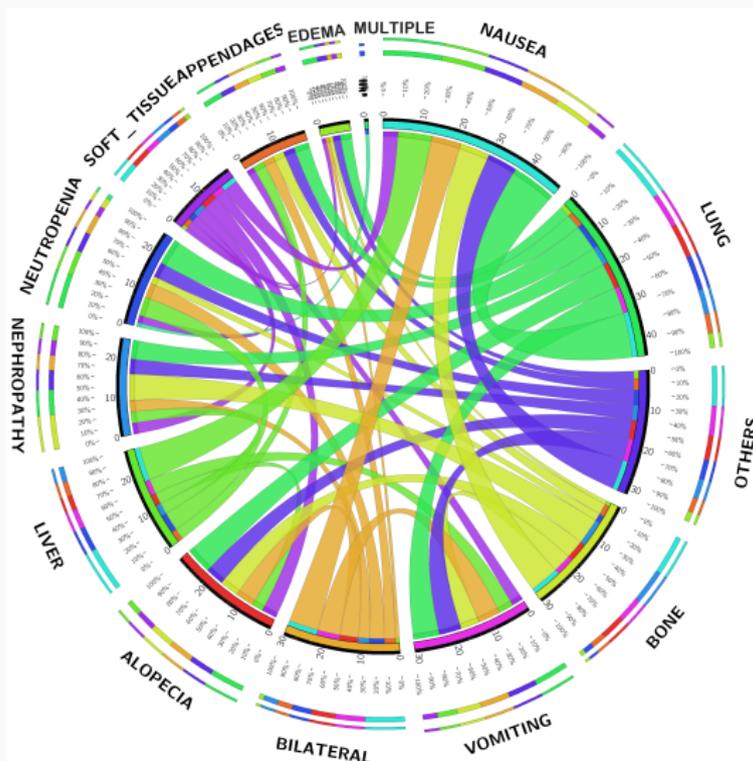


Figure: Symptoms list vs metastasis location in a circular layout

VISUALIZING DATA STRUCTURE IN A CIRCULAR LAYOUT

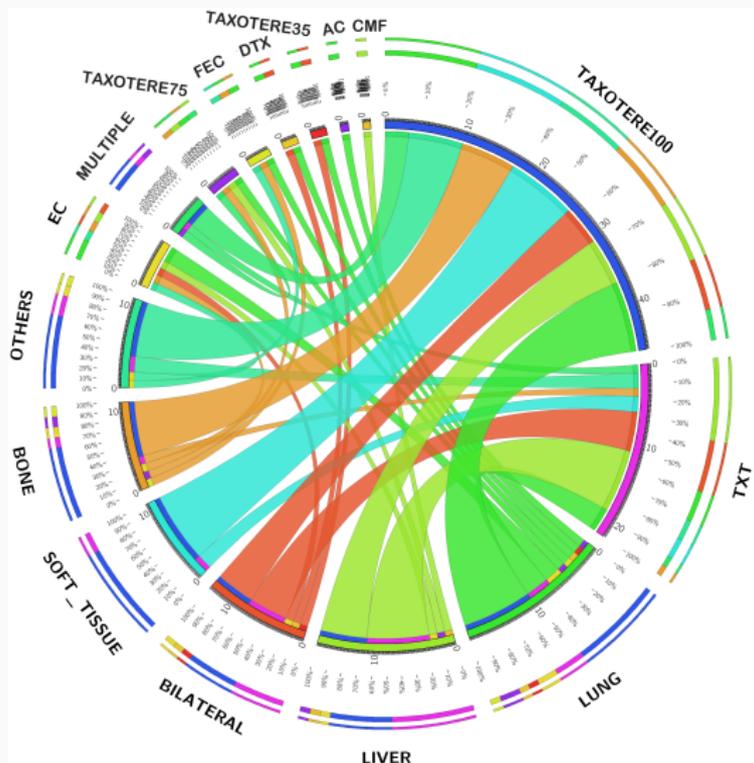


Figure: Metastasis location vs recommended treatment in a circular layout

VISUALIZING DATA STRUCTURE IN A CIRCULAR LAYOUT

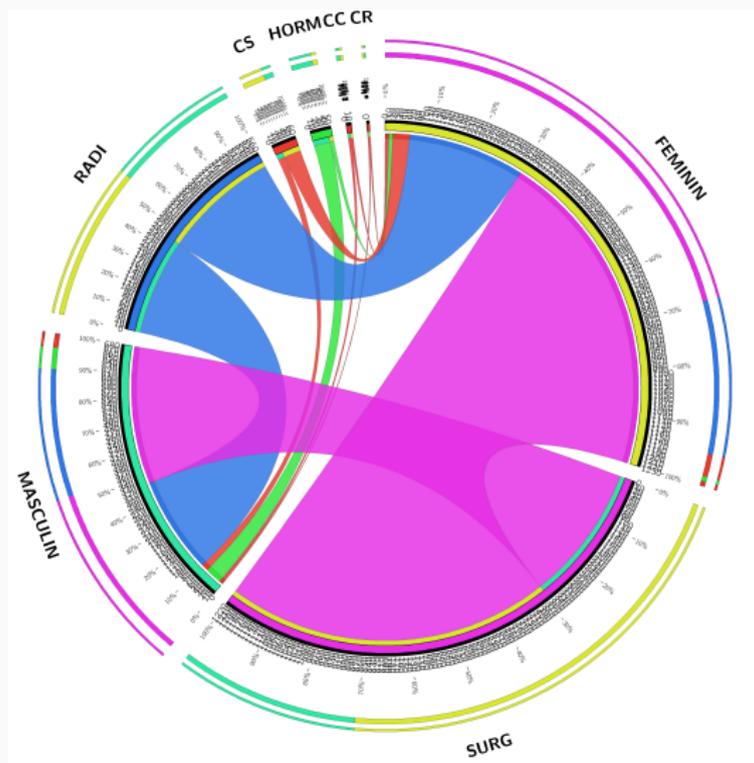


Figure: Applied treatments vs patient sex

CONCLUSIONS AND FUTURE WORK

- there is a real potential to improve EHR systems by using Conceptual Knowledge Processing tools;
- FCA is a valuable candidate in the search for a versatile, portable and very intuitive knowledge discovery, processing, and representation platform;
- invitation to both researchers and clinicians to think on and discuss the usability of FCA in medical health care data environments;
- Future discussions may also include some FCA varieties or extensions (like pattern structures or polyadic FCA).

Special thanks to Dr. Daniela Grecea, medical investigator and her team.

This paper is a result of a doctoral research made possible by the financial support of the Sectoral Operational Programme for Human Resources Development 2007-2013, co-financed by the European Social Fund, under the project POSDRU/187/1.5/S/155383-“Quality, excellence, transnational mobility in doctoral research”.

Questions?

ANALYSING THE EFFECT OF CHANGING THE EDUCATIONAL METHODS BY USING FORMAL CONCEPT ANALYSIS

SoftCOM 2016

Sanda Dragoș, Diana Haliță, Christian Săcărea

September 22, 2016

Babeș-Bolyai University, Cluj-Napoca, România

Motivation

Overview of the Analysis

Data Preparation

Data Mining by using FCA

Conclusions and Future Work

Motivation

Overview of the Analysis

Data Preparation

Data Mining by using FCA

Conclusions and Future Work

Motivation

Overview of the Analysis

Data Preparation

Data Mining by using FCA

Conclusions and Future Work

Motivation

Overview of the Analysis

Data Preparation

Data Mining by using FCA

Conclusions and Future Work

Motivation

Overview of the Analysis

Data Preparation

Data Mining by using FCA

Conclusions and Future Work

MOTIVATION

- online educational platforms are an important part of the education system;
- web analytics are not precise enough for educational content;
- new techniques: Formal Concept Analysis;
- 'who is doing what' in an e-learning platform?

OVERVIEW OF THE ANALYSIS

Investigating users behavior may help educators:

- users behavioral patterns vs. educational performance;
- a comparative study of behavioral patterns in the use of PULSE, between two different generations of students;

Interaction with the web system and the learning resources

- to evaluate on-line activities and determine how to improve the e-learning instrument used;
- in finding the most effective structure of presenting the teaching material.

But, it can also help to:

- discover student needs;
- discover how students learn;
- if and how are the students involved on-line or in class.

DATA PREPARATION

- data cleaning;
- time interval granularity: weeks;
- HTTP sessions;
- login ID;
- records about student performance and attendance;
- comparative view on two time periods, i.e., the second semesters of the academic years 2012-2013 and 2014-2015;
- subject: Operating Systems;
- 25 registered students (2012-2013), 59 registered students(2014-2015).

Visited pages ⇒ classes of access files

- after the login phase: HOME;
- laboratory content: LAB;
- lecture and teaching support: LECTURE;
- other classes: FAQ, FEEDBACK, NEWS, LOGOUT, CHANGE;

Internal Referrers ⇒ classes of access files

- after the login phase: HOME;
- laboratory content: LAB;
- lecture and teaching support: LECTURE;
- other classes: FAQ, FEEDBACK, NEWS, LOGOUT, CHANGE;

External Referrers ⇒ classes of access files

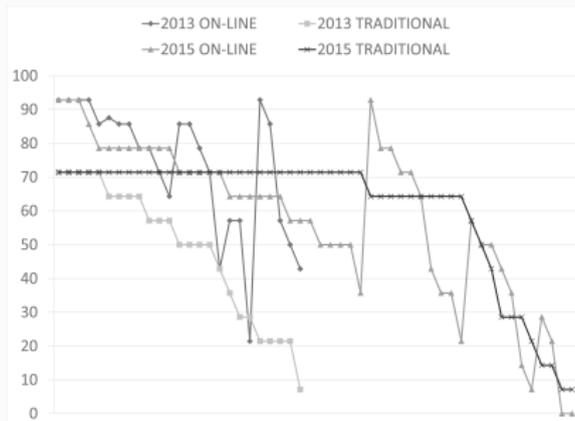
- direct accesses: DIRECT;
- facebook, google, mail: SOCIAL.

DATA MINING BY USING FCA

DATA MINING BY USING FCA

- **Attendance:** how actively involved were the students both on-line and in class;
- **Observe:** offline behavior \neq online behavior;

	2013	2015
online	73.5%	60.86%
offline	48.57%	63.14%



WHO IS DOING WHAT ON PULSE?

- what are students mostly accessing (i.e., which access classes) and from where?
- how are students handing the assignments?

WHO IS DOING WHAT ON PULSE?

Access File Classes visited in each academic year

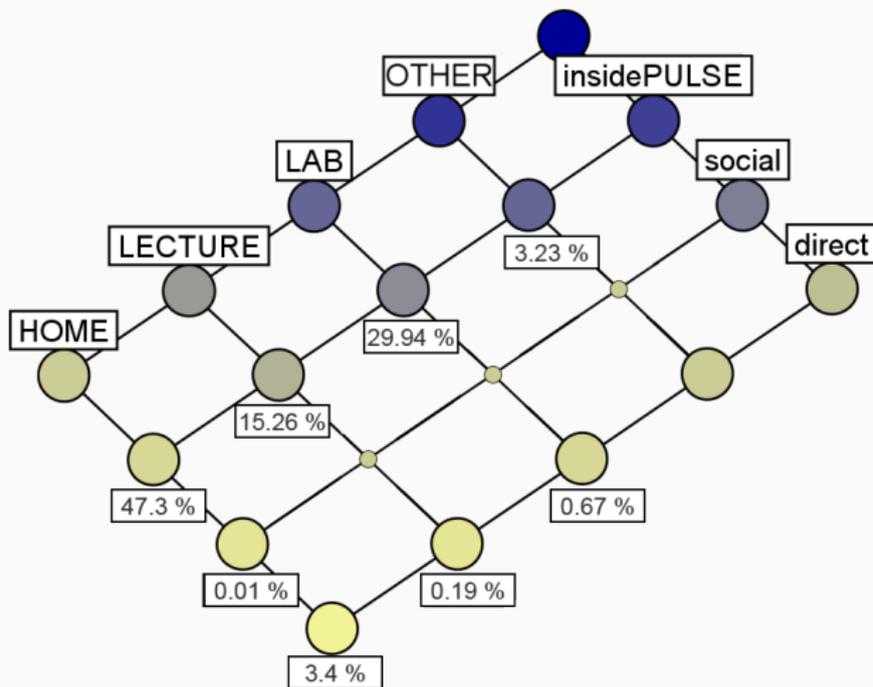


Figure: 2013

WHO IS DOING WHAT ON PULSE?

Access File Classes visited in each academic year

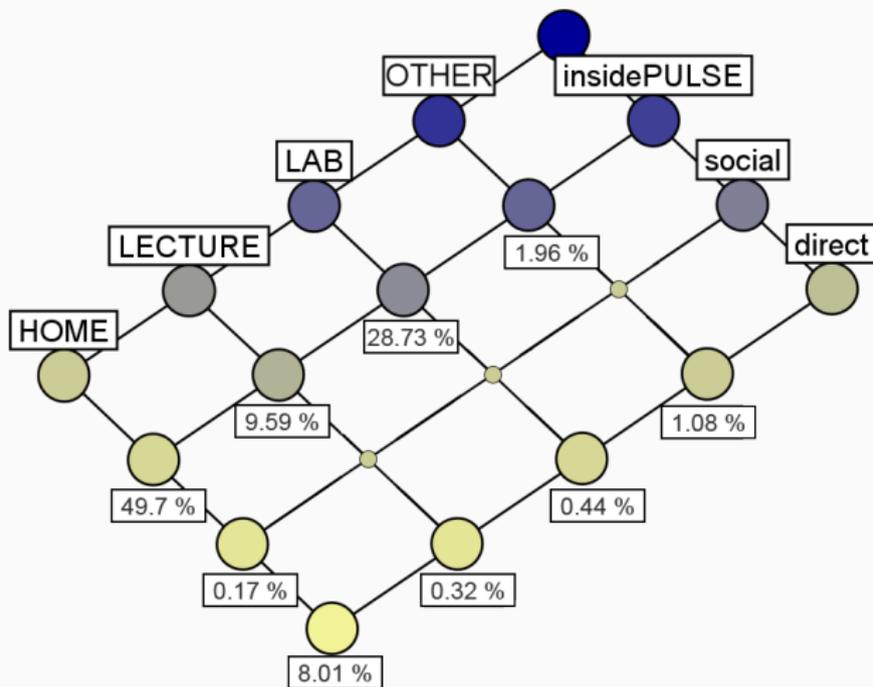


Figure: 2015

WHO IS DOING WHAT ON PULSE?

Handing lab assignments - dyadic view

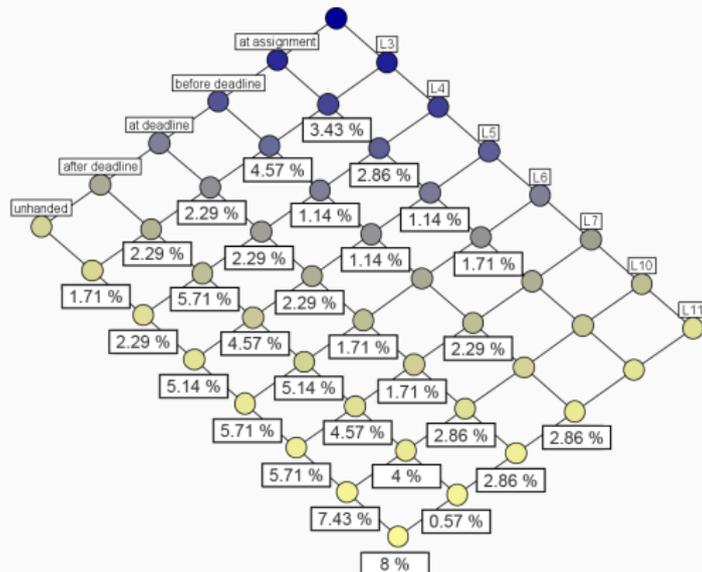


Figure: 2013

WHO IS DOING WHAT ON PULSE?

Handing lab assignments - dyadic view

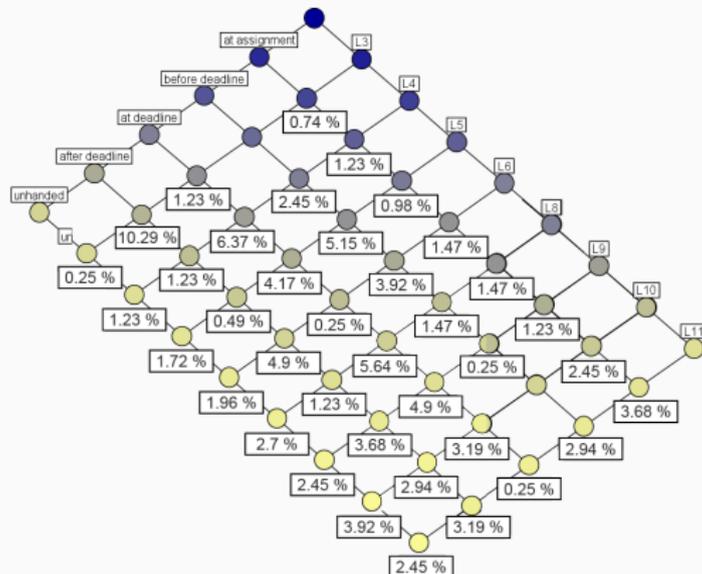


Figure: 2015

WHO IS DOING WHAT ON PULSE?

Handing lab assignments for Lab10 - circular view of triadic data

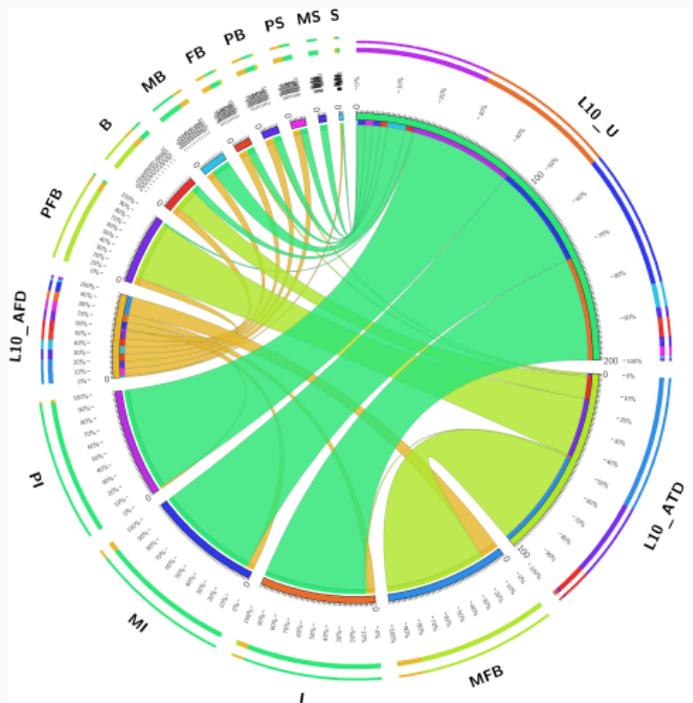


Figure: 2013

WHO IS DOING WHAT ON PULSE?

Handing lab assignments for Lab10 - circular view of triadic data

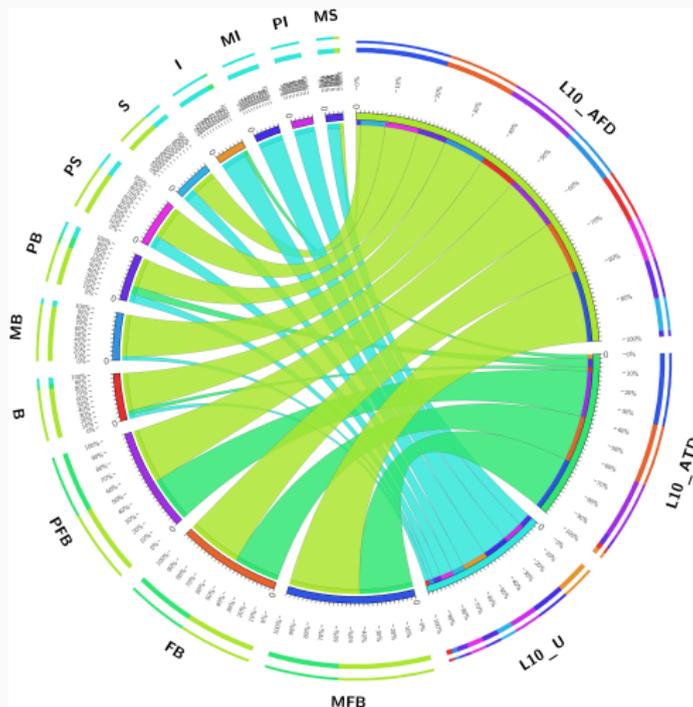


Figure: 2015

WHO IS DOING WHAT ON PULSE?

Correlation between online behavior and educational performance

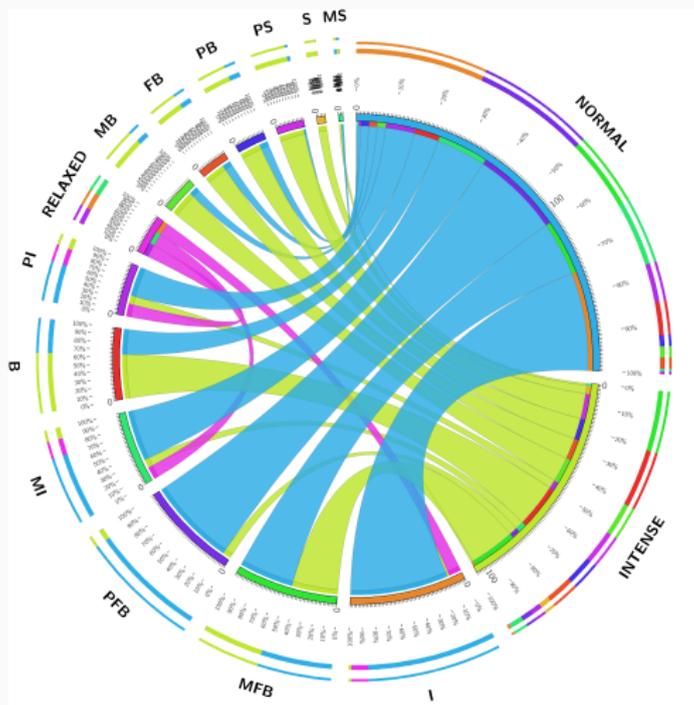


Figure: 2013

WHO IS DOING WHAT ON PULSE?

Correlation between online behavior and educational performance

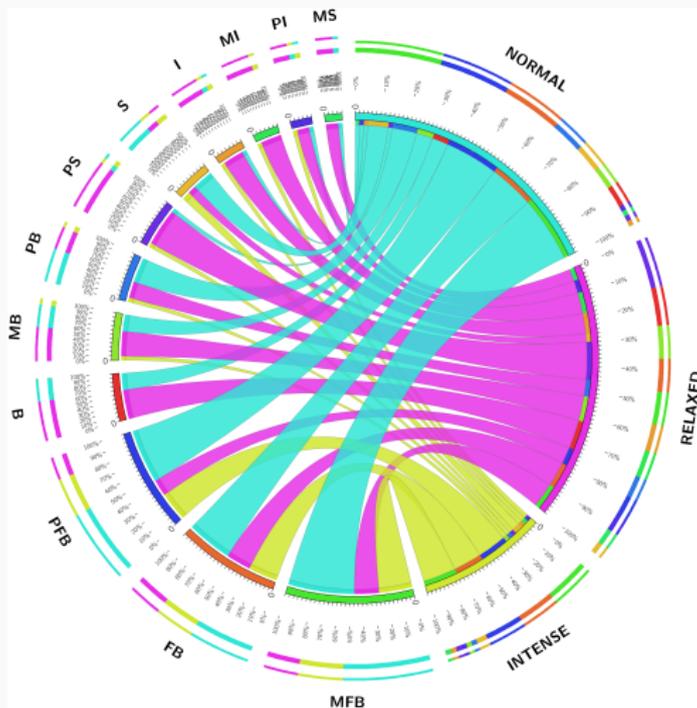


Figure: 2015

CONCLUSIONS AND FUTURE WORK

- comparative study on two different sets of data about students academic performance versus two different educational approaches;
- the developed methods are not restricted to PULSE or to an e-learning environment;
- detecting behavioral patterns and extracting, processing and representing knowledge with FCA tools is efficient and helpful for any knowledge management task.

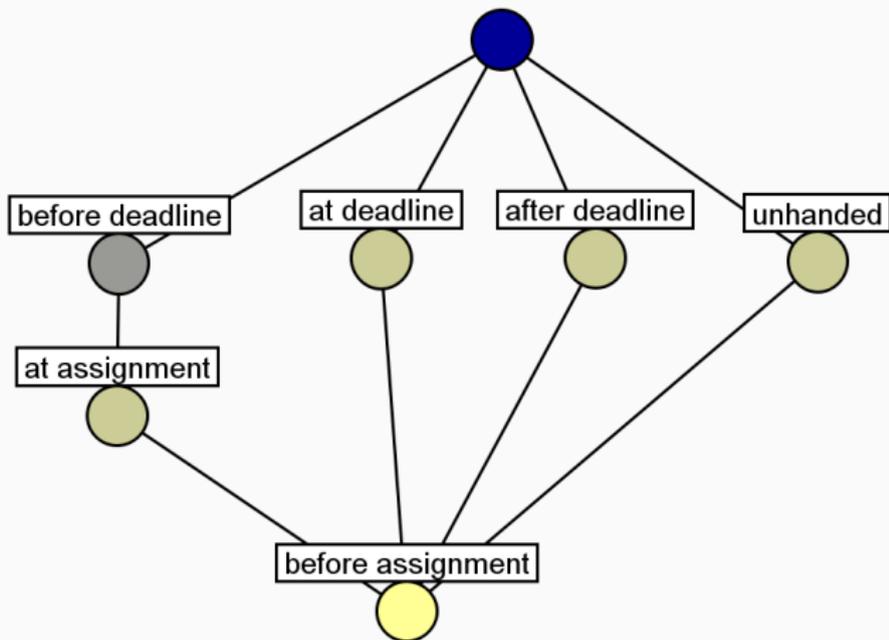


Figure: Early Student

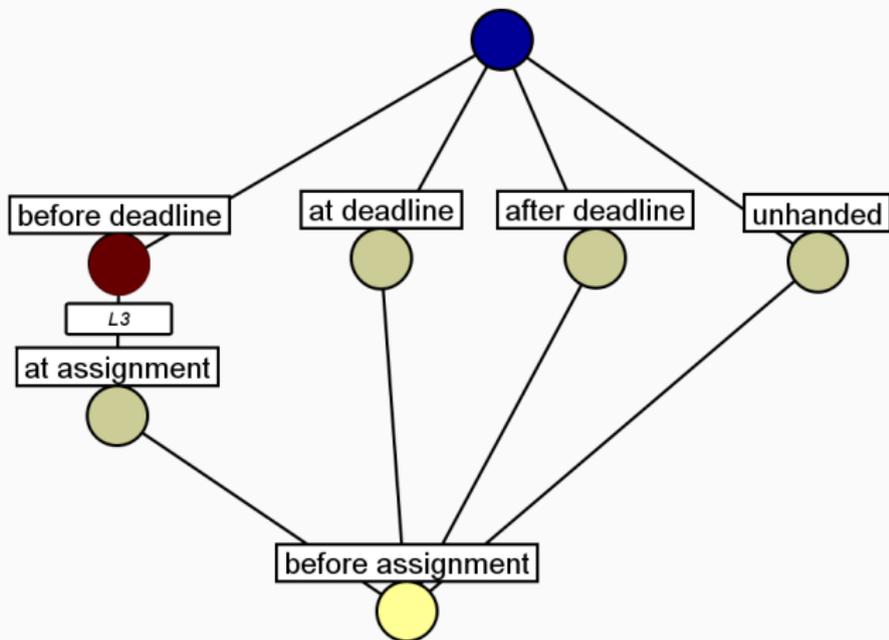


Figure: Early Student

FUTURE WORK - DISCOVER USER TRAILS

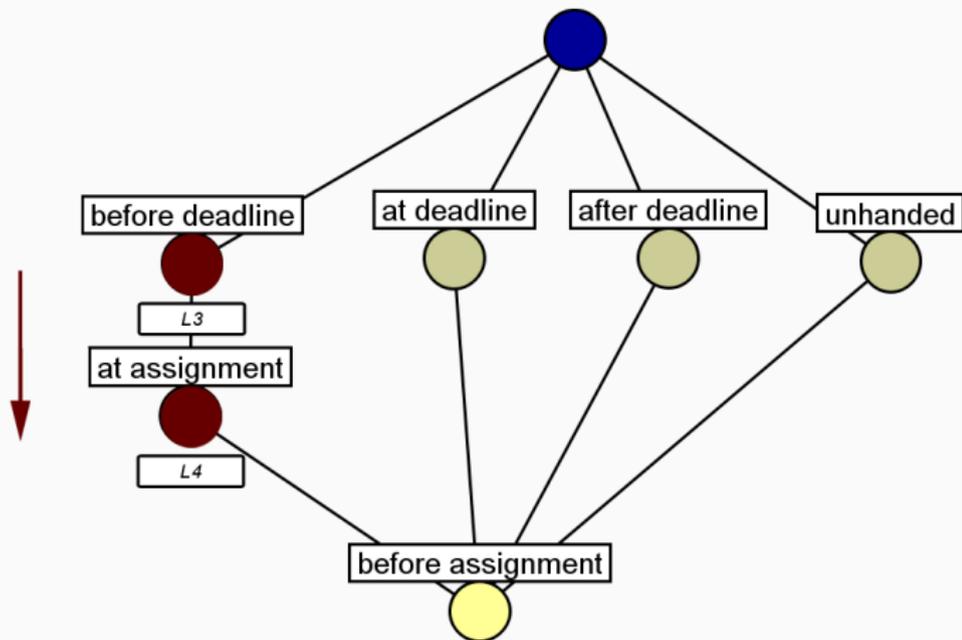


Figure: Early Student

FUTURE WORK - DISCOVER USER TRAILS

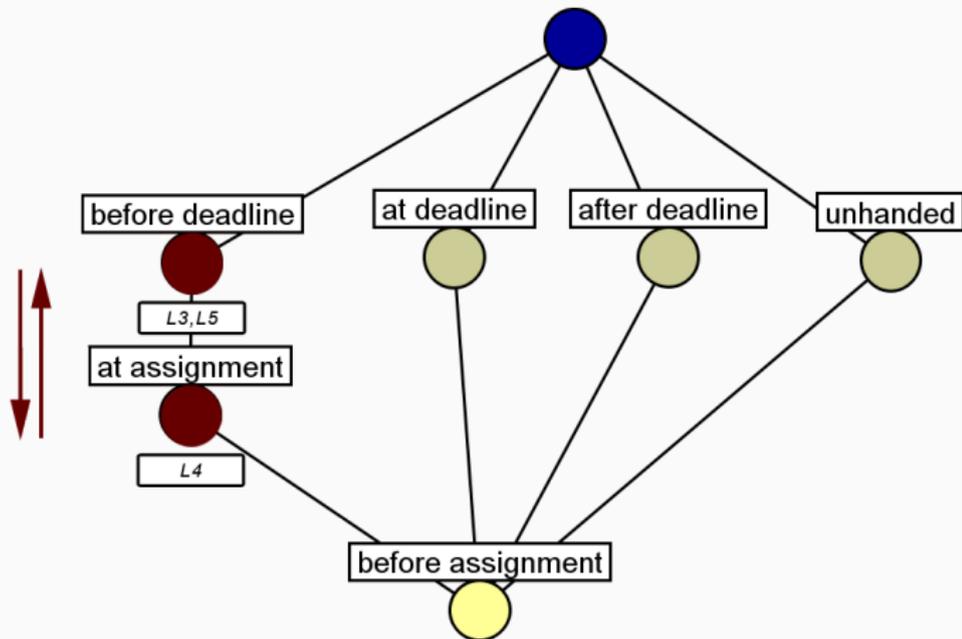


Figure: Early Student

FUTURE WORK - DISCOVER USER TRAILS

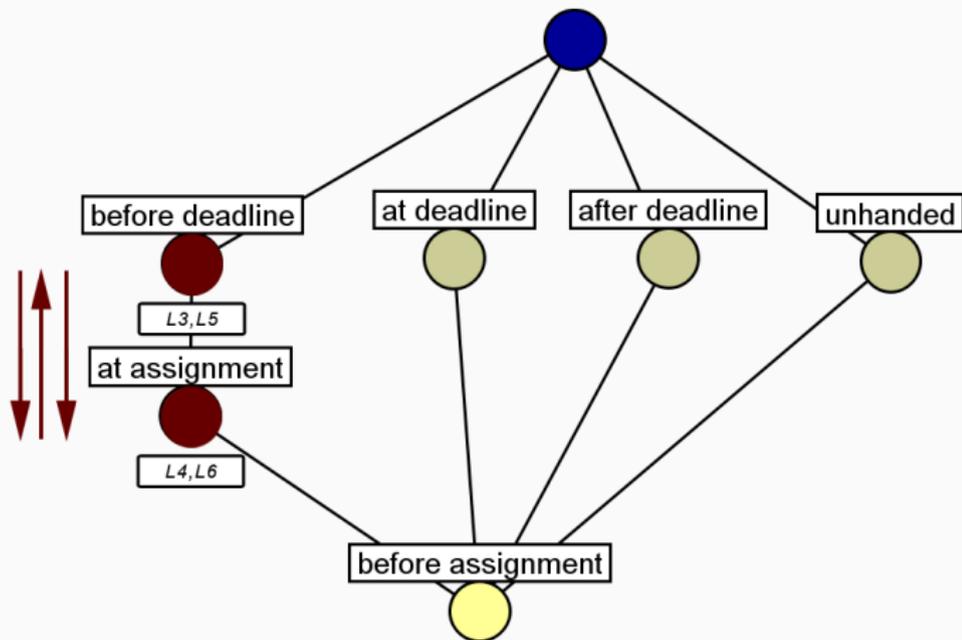


Figure: Early Student

FUTURE WORK - DISCOVER USER TRAILS

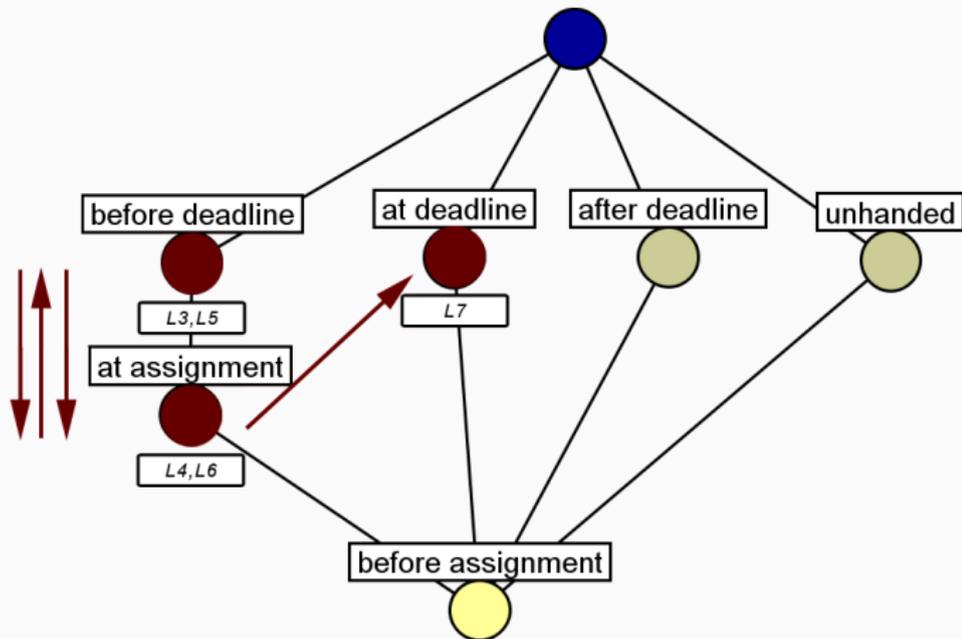


Figure: Early Student

FUTURE WORK - DISCOVER USER TRAILS

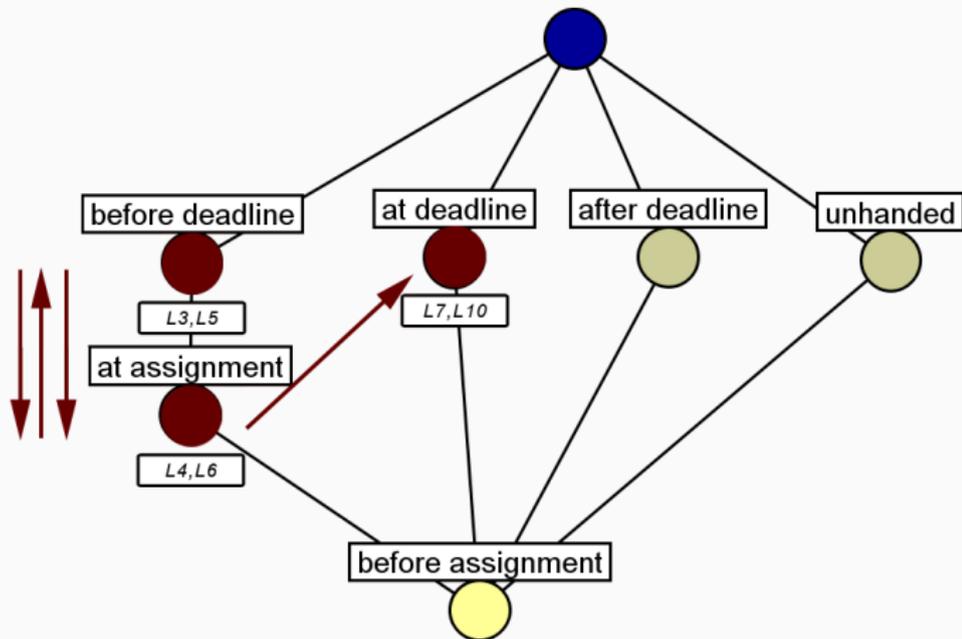


Figure: Early Student

FUTURE WORK - DISCOVER USER TRAILS

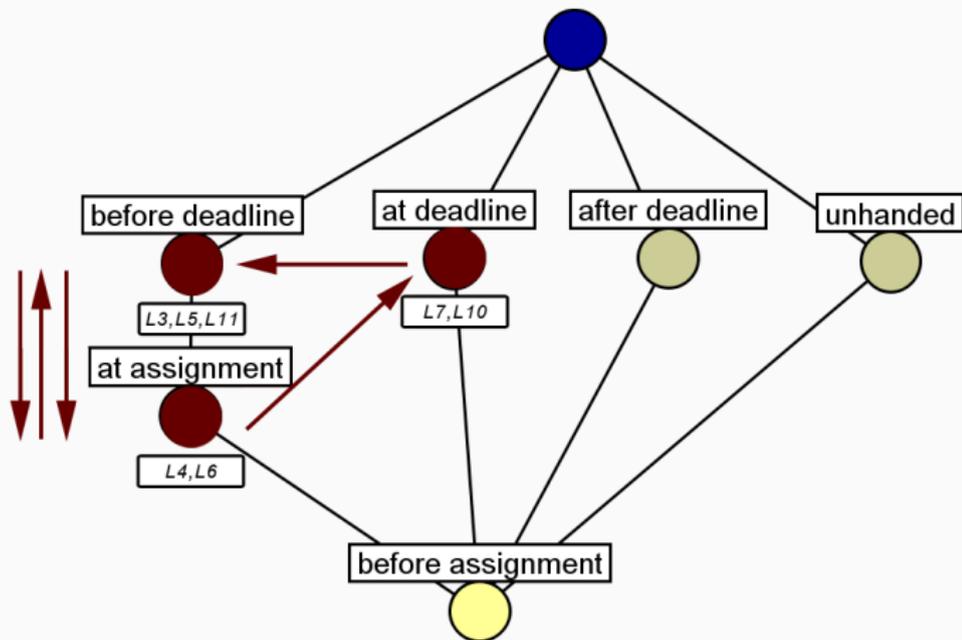


Figure: Early Student

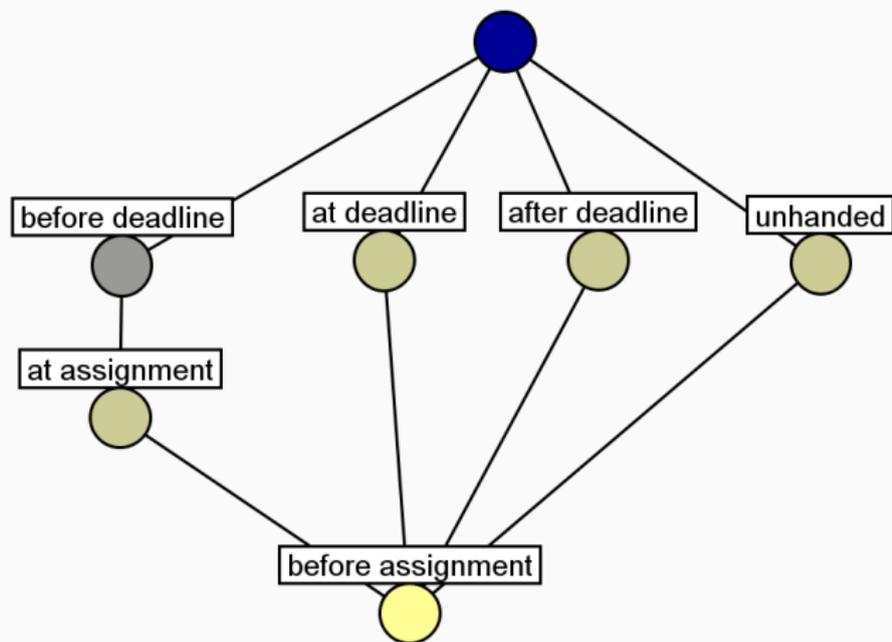


Figure: Common Student

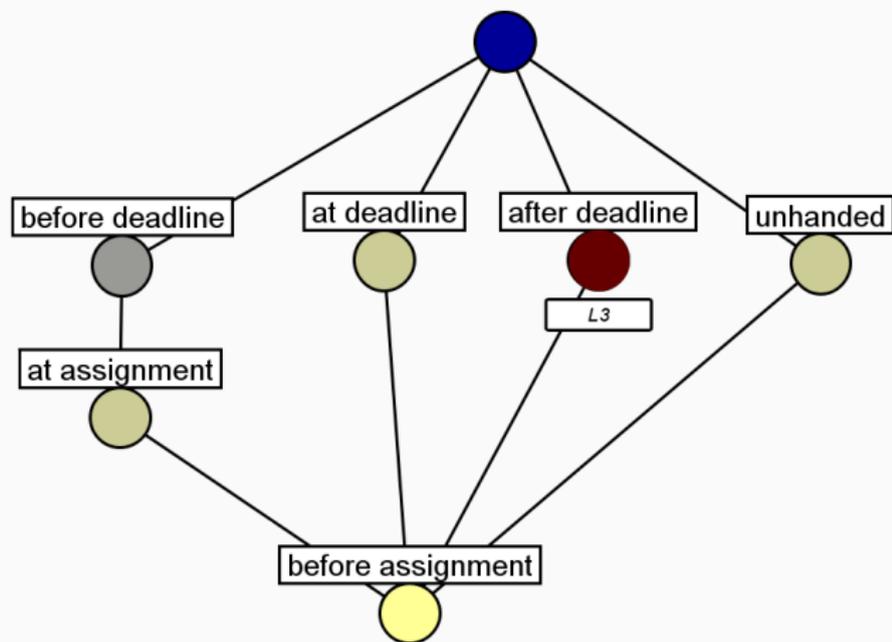


Figure: Common Student

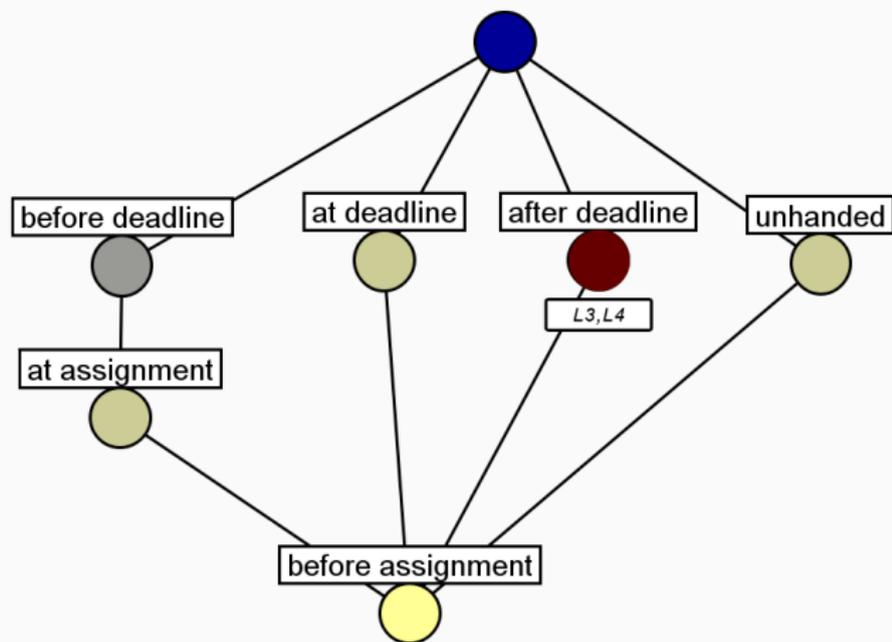


Figure: Common Student

FUTURE WORK - DISCOVER USER TRAILS

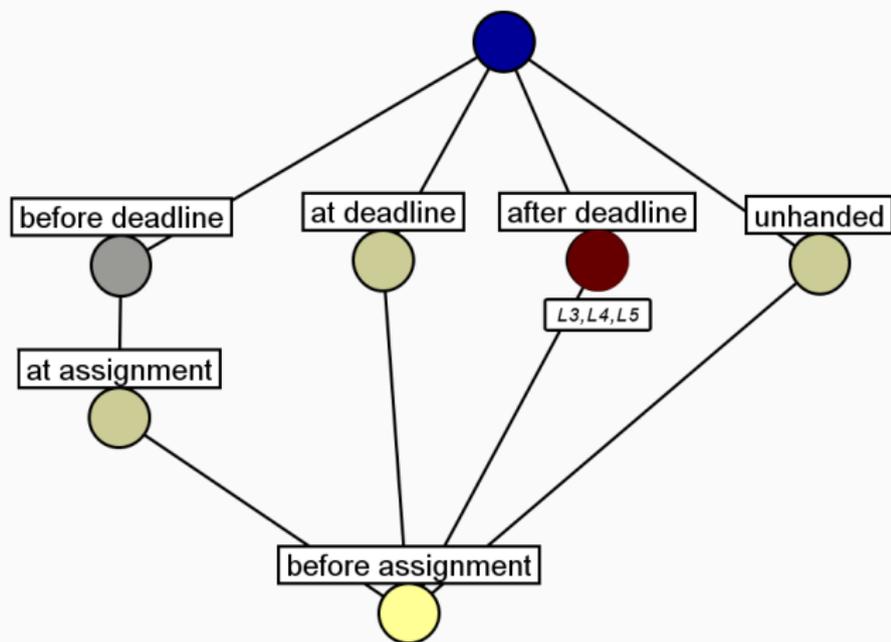


Figure: Common Student

FUTURE WORK - DISCOVER USER TRAILS

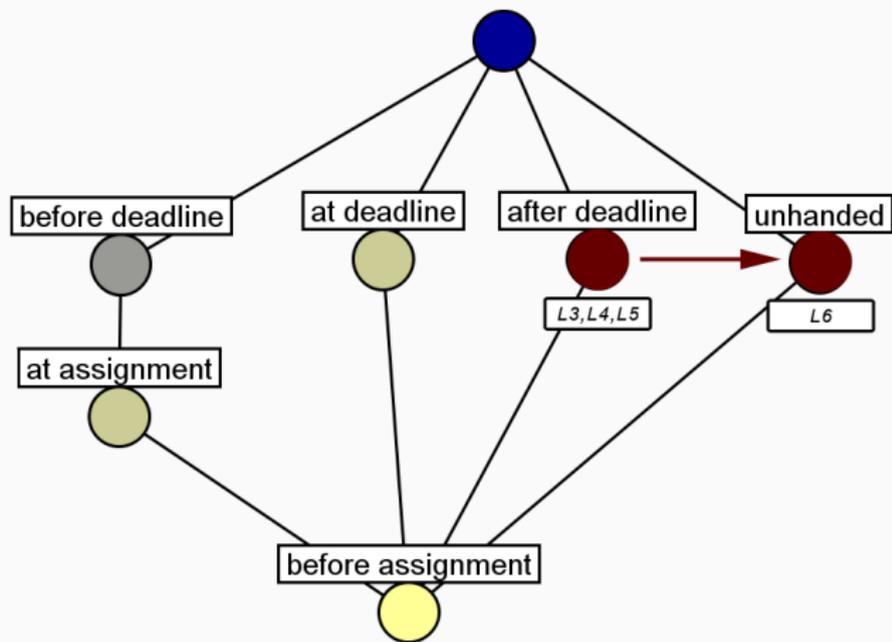


Figure: Common Student

FUTURE WORK - DISCOVER USER TRAILS

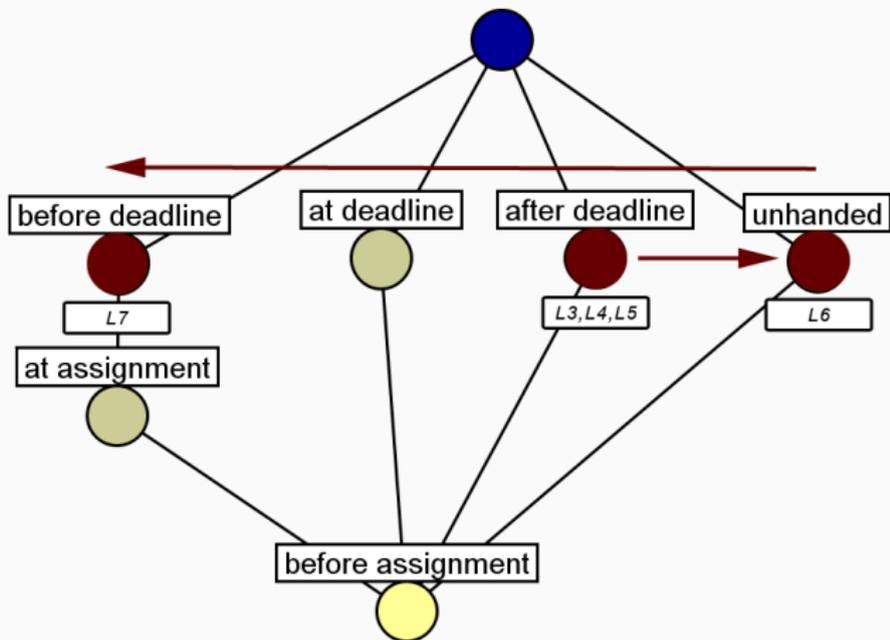


Figure: Common Student

FUTURE WORK - DISCOVER USER TRAILS

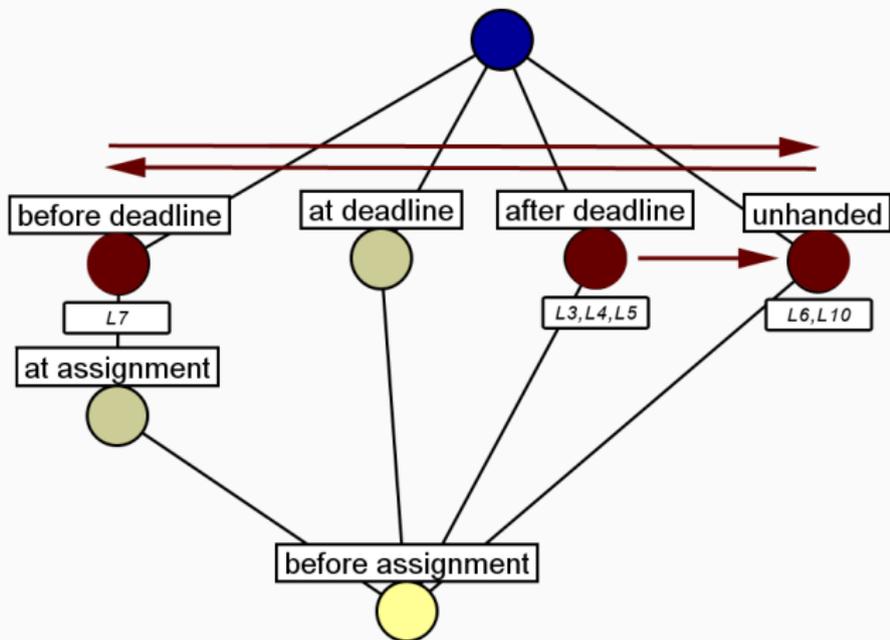


Figure: Common Student

FUTURE WORK - DISCOVER USER TRAILS

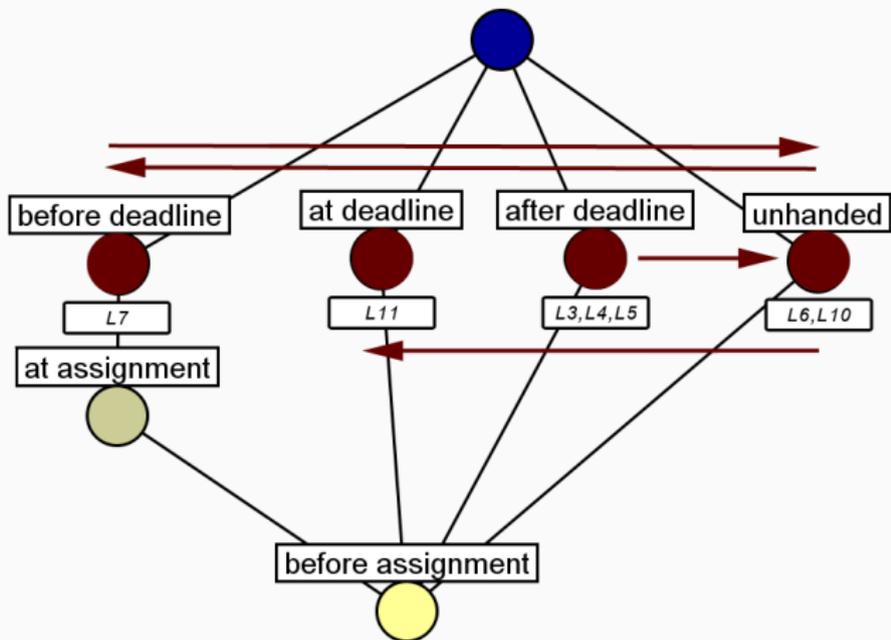


Figure: Common Student

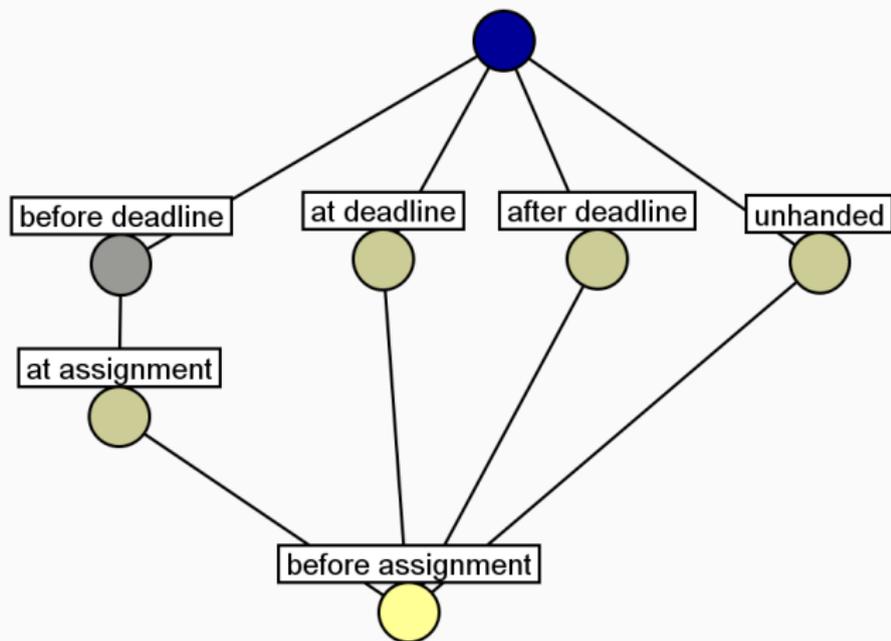


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

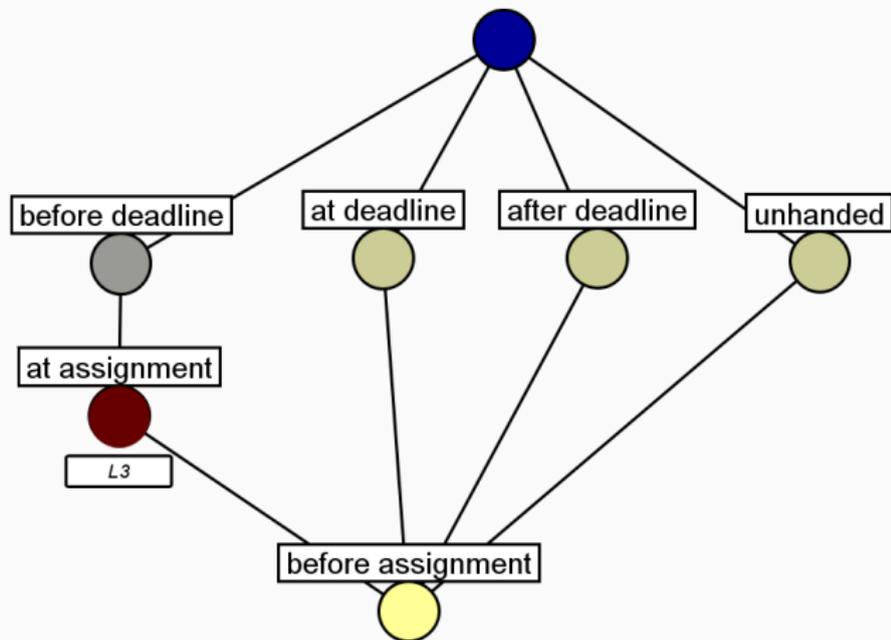


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

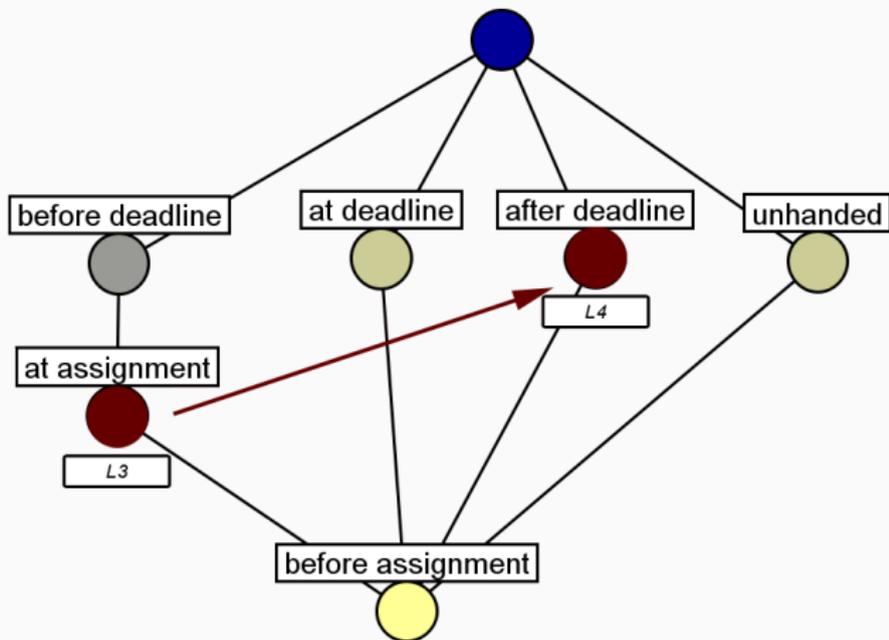


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

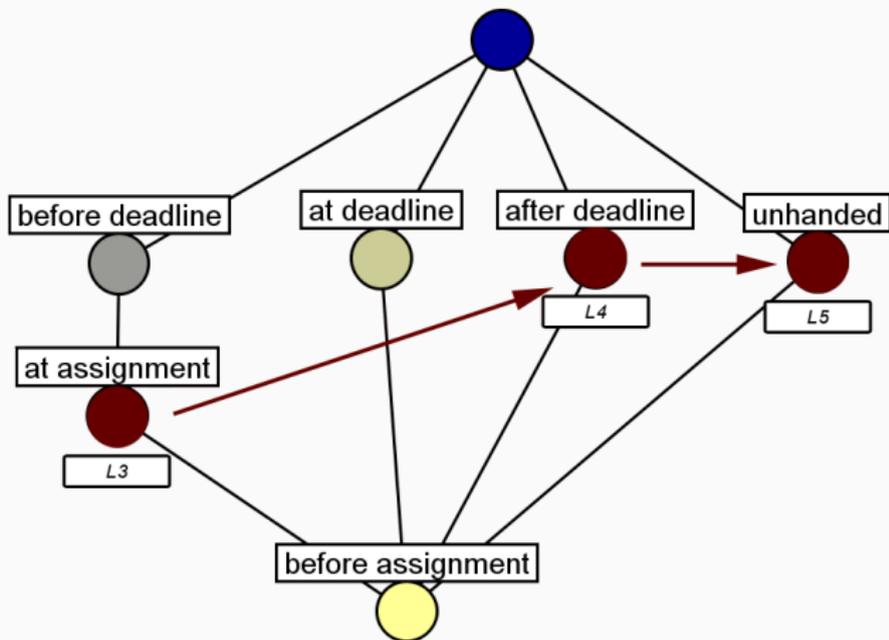


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

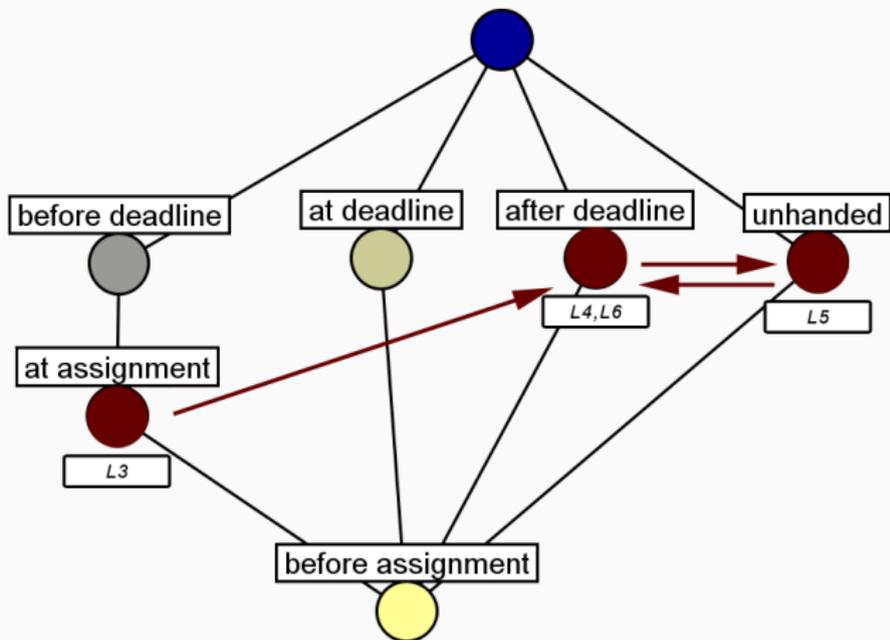


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

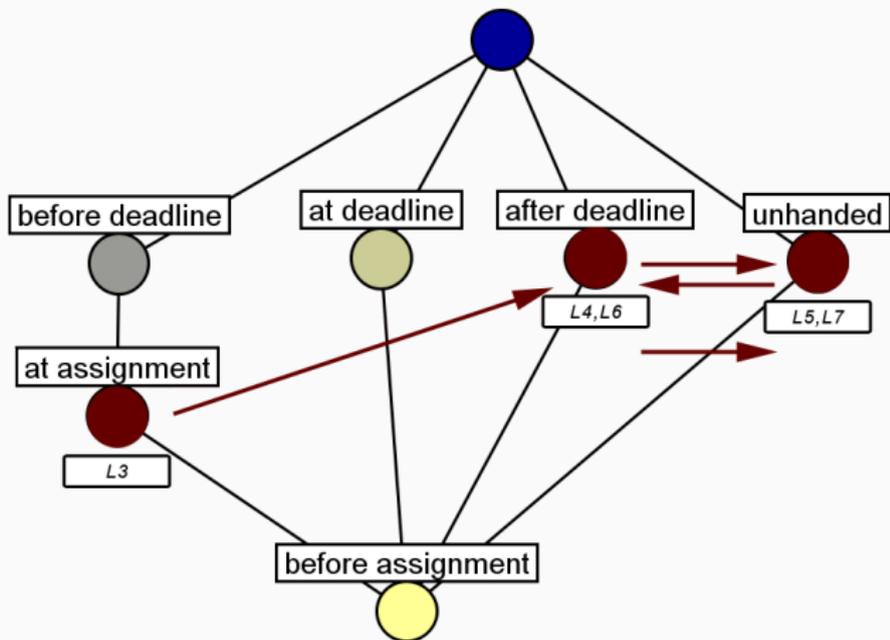


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

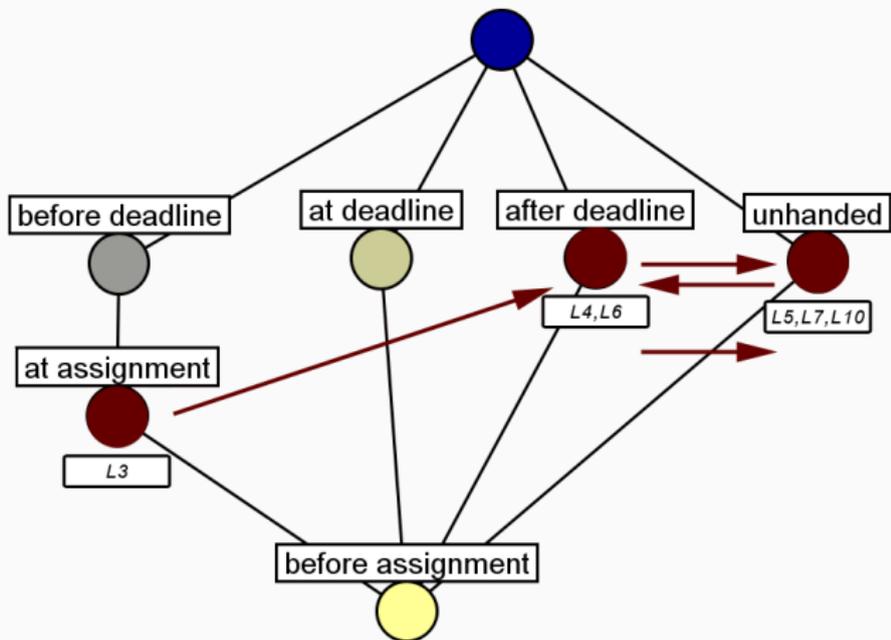


Figure: Late Rise Student

FUTURE WORK - DISCOVER USER TRAILS

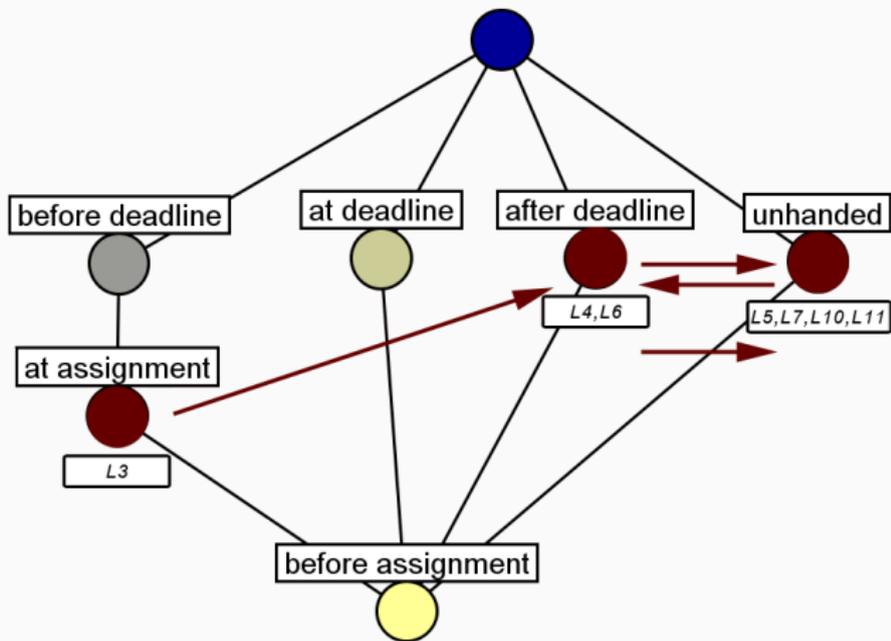


Figure: Late Rise Student

This paper is a result of a doctoral research made possible by the financial support of the Sectoral Operational Programme for Human Resources Development 2007-2013, co-financed by the European Social Fund, under the project POSDRU/187/1.5/S/155383-“Quality, excellence, transnational mobility in doctoral research”.

Questions?

Thank You for Your
Attention!