

# Distilling Conceptual Structures from Weblog Data Using Polyadic FCA

**Sanda Dragoș, Diana Haliță and Christian Săcărea**

Babeș-Bolyai University, Department of Computer Science, Cluj-Napoca,  
România

[sanda@cs.ubbcluj.ro](mailto:sanda@cs.ubbcluj.ro), [diana.halita@ubbcluj.ro](mailto:diana.halita@ubbcluj.ro), [csacarea@math.ubbcluj.ro](mailto:csacarea@math.ubbcluj.ro)

July 2016



# PREREQUISITES

## What we want

- observe and understand student behavior and the interaction between them;
- determine students with similar on-line access behavior (i.e., repetitive browsing habits);
- find bundle of users with similar interests;
- identify common sub-behaviors of more complex behaviors.



# PREREQUISITES

## Tools

- Cosine similarity measure;
- knowledge is extracted from the data set for knowledge discovery using the classical FCA approach;
- the ToscanaJ knowledge management suite enabled a systematic exploration of the data set in order to detect behavioral patterns.

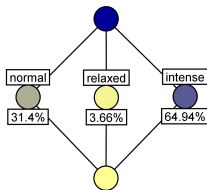


# PREREQUISITES

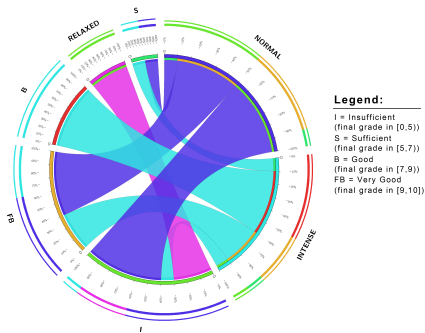
## Data preprocessing: clusters

- visit chains;
- academic week time intervals;
- 10 classes of access files:
  - information about the lecture and the laboratories (I);
  - information about the practical examination (PE);
  - information about the written examination (WE);
  - theoretical support for all laboratories (LT);
  - all laboratory assignments (LA);
  - all laboratory examples (LE);
  - slides and notes for all lecture (Ls));
  - overview information about the lectures (L);
  - details on all lecture test papers (LPs);
  - overview on the test papers given during lectures (LP).

# USER BEHAVIOR - INVESTIGATIONS TILL NOW



(a) Distribution over 3 classes



(b) Correlation: grades and behavior

**Figure:** Users quantitative behavioral patterns

# USER BEHAVIOR - INVESTIGATIONS TILL NOW

## Qualitative behavioral patterns: Attractors

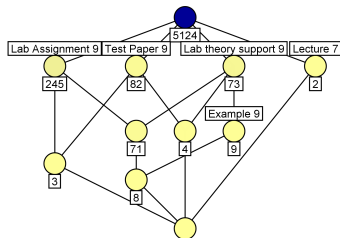
Attractors are behavioral patterns to which users adhere while using the educational system, and are distilled from frequently visited chain of pages where some deviation from the visit habit is allowed.

- 1 Educational attractors;
- 2 Popular attractors;
- 3 Critical attractors.

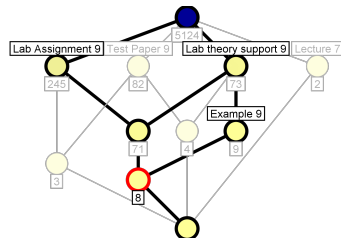


# EDUCATIONAL ATTRACTORS

- reflect the educational purpose of the instructor and convince or persuade users to adhere to them;



(a) Related teaching support



(b) Laboratory support (assignment, theory, example)

Figure: Educational attractors for the 9th Laboratory

# USER BEHAVIOR - CURRENT INVESTIGATIONS

## Building the context

We determined pairs of chains that have a similarity of at least 80%:

|               |       |            |       |            |
|---------------|-------|------------|-------|------------|
| similar chain | user1 | week user1 | user2 | week user2 |
|---------------|-------|------------|-------|------------|

- 1 *objects* - chains;
- 2 *attributes* - weeks for user1;
- 3 *conditions* - weeks for user2.

# THE TRIADIC APPROACH

For student  $K$  there are a list of 20 non-trivial triconcepts. We start with:  $([L - LPs, L - LPs - WE], [w1 = 15, w1 = 16], [w2 = 16])$

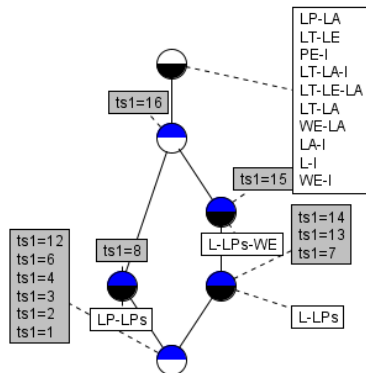


Figure: Navigation along perspective 3, condition  $[w2=16]$

# THE TRIADIC APPROACH

$([L - LPs, L - LPs - WE], [w1 = 15, w1 = 16], [w2 = 16])$

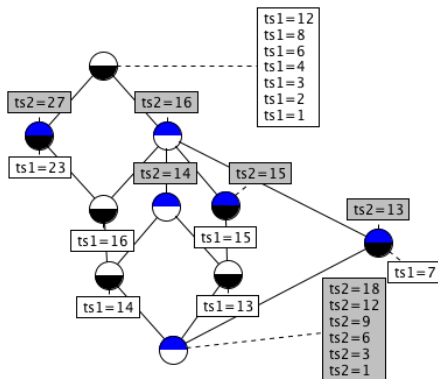


Figure: Navigation along perspective 1, chain  $[L-LPs]$

# AN EXTENSION: 4FCA

- ① *objects* - user 2;
- ② *attributes* - weeks for user1;
- ③ *conditions* - weeks for user2;
- ④ *states* - chains.

```

1 ([u2=H,K,O,T], [w1=2], [w2=2], [LA-I])
2 ([u2=E], [w1=1,w1=2,w1=3], [w2=6], [LA-I])
3 ([u2=K], [w1=2], [w2=2], [LA-I,WE-LA-I])
4 ([u2=B,H,T], [w1=3], [w2=2], [LT-LA-I])
5 ([u2=H], [w1=3], [w2=2], [LT-LA-I,LT-LA])
6 ([u2=K], [w1=3], [w2=3], [LT-LA-I,L-LA-I,LT-LA,L-LA,L-WE-LA])
7 ([u2=O], [w1=3], [w2=2], [L-LA-I])
8 ([u2=B], [w1=9], [w2=7], [L-PE-LA-I])
9 ([u2=K], [w1=16], [w2=16], [LP-LPs-PE-LA-I,LP-LPs,L-LPs,L-WE,Ls-L-LP])

```

Figure: Tetradic analysis for user *K* for the chain "LA-I"

# QUANTITATIVE MEASUREMENTS ON 4FCA RESULTS

We selected two most active users:  $K$  and  $H$ .

|   | meta-<br>class | no. of concepts |                                         |       |     |     |     |     |     |                |
|---|----------------|-----------------|-----------------------------------------|-------|-----|-----|-----|-----|-----|----------------|
|   |                | tot.            | behavior similar with sets of $n$ users |       |     |     |     |     |     | self<br>behav. |
|   |                |                 | n=9                                     | n=6,7 | n=5 | n=4 | n=3 | n=2 | n=1 |                |
| K | 15             | 234             | 1                                       | 0     | 1   | 10  | 22  | 72  | 162 | 58             |
| H | 13             | 266             | 1                                       | 2     | 6   | 13  | 22  | 41  | 179 | 26             |

**Table:** Comparison results on users  $K$  and  $H$



# QUANTITATIVE MEASUREMENTS ON 4FCA RESULTS

| Class                                  | K         | H          | Class len |
|----------------------------------------|-----------|------------|-----------|
| <b>Lab Theory, Lab Assignment</b>      | <b>51</b> | <b>104</b> | <b>2</b>  |
| Lab Assignment, Info                   | 15        | 0          | 2         |
| Lecture, Info                          | 18        | 0          | 2         |
| Lecture, Lab Theory                    | 0         | 18         | 2         |
| Lecture, Lab Theory, Lab Assignment    | 0         | 9          | 3         |
| <b>Lecture, Lab Assignment</b>         | <b>15</b> | <b>13</b>  | <b>2</b>  |
| <b>Lecture, Lecture Paper</b>          | <b>6</b>  | <b>35</b>  | <b>2</b>  |
| Lecture, Lecture Paper, Lab Assignment | 0         | 9          | 3         |
| <b>Lecture, Lecture Papers</b>         | <b>62</b> | <b>16</b>  | <b>2</b>  |
| <b>Lecture Paper, Lecture Papers</b>   | <b>31</b> | <b>17</b>  | <b>2</b>  |
| <b>Lecture Papers, Lab Assignment</b>  | <b>14</b> | <b>9</b>   | <b>2</b>  |
| <b>Lecture Papers, Written Exam</b>    | <b>13</b> | <b>6</b>   | <b>2</b>  |
| <b>Practical Exam, Info</b>            | <b>36</b> | <b>45</b>  | <b>2</b>  |

**Table:** Top classes for students *K* and student *H*.



# BUNDLE OF USERS

```

1  For student K
2    users with similar behavior = G, H, K
3    weeks of behavioural occurrence for K = 7, 13
4    weeks of behavioural occurrence for users with similar behavior with
      K = 5, 6, 7, 11, 14
5
6  For student H
7    users with similar behavior = B, D, F, G, H, I, W, K, L, M, O, Q, T,
      X
8    weeks of behavioural occurrence for H = 4, 5, 6, 7, 9
9    weeks of behavioural occurrence for users with similar behavior with
      H = 3, 4, 5, 6, 7, 8, 9, 11, 12, 14, 15, 17, 20

```

Figure: Stats for students  $K$  and  $H$  on chain “ $L$ - $LP$ ”

# CONCLUSIONS AND FUTURE WORK

## Conclusions

- use n-adic FCA to determine repetitive behavioral patterns;
- determined similar behavior that fit into classes of our interest;
- determined which are the meta-classes that users use;
- determined bundle of users with similar behavior;
- understand how students are using the e-learning platform.

## Future Work

- use 5-adic FCA in order to determine trend-setters.



Thank you for your attention!  
Questions?

