EduVis: Visualizing Educational Information

Sandra Gama
INESC-ID and Instituto
Superior Técnico
Av. Rovisco Pais, 1
Lisboa, Portugal
sandra.gama@ist.utl.pt

Vilma Jordão
INESC-ID and Instituto
Superior Técnico
Av. Rovisco Pais, 1
Lisboa, Portugal
vilma.jordao@ist.utl.pt

Daniel Gonçalves
INESC-ID and Instituto
Superior Técnico
Av. Rovisco Pais, 1
Lisboa, Portugal
daniel.goncalves@inesc-id.pt

Abstract
A successful analysis of educational processes may help
enhance success. Data Mining techniques, despite
allowing analysis of such data, result in an extensive set of
symbolic patterns that are difficult to understand.
Visualization may overcome this limitation due to its
potential to display large quantities of data while
alleviating cognitive load. We developed a visualization
that allows the analysis of patterns obtained by using
educational data mining techniques to gather patterns of
interdependences among courses in a university program.
We created EduVis, a coordinated visualization which
takes advantage of two different, complementary, tools: a
multi-layered visualization and a multi-matrix
representation of courses and corresponding relationships.
Preliminary user tests have shown that EduVis makes
important patterns immediately perceivable, suggesting
that a small number of adjustments will realize its full
potential for visualizing educational information.

Author Keywords
Educational Data Representation, Information
Visualization, Human-Machine Interfaces

ACM Classification Keywords
H.5.2 [Computer Graphics]: Information interfaces and
presentation - User Interfaces.

Introduction

The number of students has grown over the last decades, in universities [1] and online education, with the profusion of MOOC (Massive Open Online Courses). A large set of data emerges from these processes which, if explored effectively, may help improve education. EDM (Educational Data Mining) is thus becoming popular. However, results consist of an extensive set of behaviors described as textual patterns which are difficult to understand due to visual complexity. Allowing an effective interpretation of information is, thus, extremely relevant.

Visualization has the potential to overcome this challenge: it is an excellent means for displaying information, allowing data interpretation with low cognitive load [4]. Visualizing the results of EDM will enable the educational community to be aware of educational patterns that would otherwise remain unnoticed. We created EduVis, a visualization that displays educational patterns, making them easy to navigate and analyze. The displayed information is the result of EDM techniques that had previously been applied to analyze the interdependence of success among courses, providing us with information on the number of students that pass each course, as well as precedence and simultaneity relationships among courses. EduVis displays such patterns and allows their exploration by combining a multi-layered view, which displays visual connectors among interrelated courses, and a multi-matrix view, which allows filtering and further exploration. Hence, EduVus makes general information immediately perceivable and allows interactive manipulation for further data exploration.

In this paper we discuss related work and introduce the mining patterns we have used. We then describe EduVis, drawing conclusions and directions for future work.

Related Work

Several tools have been created for representing educational information. Regarding online education, CourseVis [3] attempts at displaying CMS data using (i) a forum representation, in which threads are depicted as spheres with size proportional to the number of students; (ii) a Cognitive Matrix, associating students' names with course concepts through a red(failure)-green(success) color scale; (ii) a matrix-like tool that shows students' data such as attendance and progress. Difficulties in graphic reading led to the creation of GISMO [2], which emphasizes course accesses and resources, allowing interactive exploration of student behavior. Concerning traditional education, AVOJ [6] provides the means for grouping students according to aspects such as grades and study habits and allows comparison through a bar chart tool with real-time statistics. Another tool has been created which allows the visualization of repeating patterns on failure and success [5], using a colored node and edge structure. Nodes depict events (exams or assignments), with width proportional to the number of students. Colors represent students' performance, allowing the differentiation of student groups by behavior and the selection of student categories. This made it possible for teachers to draw a set of conclusions regarding repetitive course failures or implications of a given course's grade on the success at other courses.

All aforementioned approaches present techniques for visualizing data in a particular context. Since our goal is to interrelate courses and show interdependences, we highlight the work by Wortman and Rheingans [5] which, although very interesting, does not allow course comparison or pattern highlighting. Hence, we developed EduVis, which represents interrelations among courses while attempting to address existing solutions' limitations.

Educational Patterns

We used the result of data mining that had been applied to nine years' data on a computer science program. Textual patterns have been generated, with the structure:

$$Pattern_i = (semester_1, ..., semester_N, total_{students}),$$

$$semester_j = course_1 \lor (course_1, ..., course_M).$$

For instance, $(fex, 2000)$ (*2000 students completed* fex *on the first semester*), or $((fex, am1), tc, 1000)$: (*1000 students completed* fex *and* am1 *on the first semester and* tc *on the second semester*).

These patterns provide information on course dependences that can be made evident by an effective visualization.

EduVis: Visualizing Educational Information

EduVis is a coordinated visualization of two main tools (Figure 1): (i) a multi-layer view, and (ii) a multi-matrix representation. The following sections present both tools and describe the general mechanisms of interaction.

**Figure 1:** EduVis.

Multi-Layer Visualization

In this representation, each layer depicts one semester of a study program, displaying courses as circles or, when data on failure is available, the course circle is sub-divided into two semicircles: the green circle, on the left, represents the number of successful students and the rightmost, red circle, shows the number of students who failed the course, following conventional western color codes [4], (Figure 2). This makes it possible to immediately understand the courses with high or low success rate as well as to visually compare course success and failure.
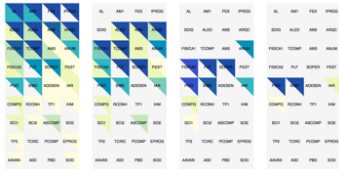
Multi-Matrix Visualization

In this view, courses are represented as matrix squares, divided into two triangles: the upper for success and the lower for failure. Before interaction (Figure 1), courses with available patterns are represented in colors within the yellow and dark blue range, while the remaning courses are depicted in gray. Brightness represents the number of patterns in which courses are involved: the lower the brightness, the more relationships with other courses.

Interaction

Interaction starts when a course is selected by moving the mouse either over a circle or semi-circle (multi-layer view), or a triangle (multi-matrix view). The multi-layer view highlights the current circle, increasing its saturation, and represents relationships of the course with others through visual connectors (cubic Bézier curves). Their thickness is proportional to the number of students who verify the behavior and Western color conventions [4] are used to assign color to curves: patterns related with success range from blue (low number of students) to green (high number); patterns that arise from failure (red semi-circles) are represented in colors ranging from yellow (low number of students) to red (high number) (Figure 2).
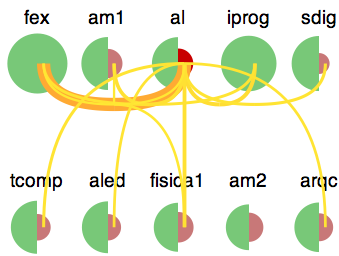
Figure 2: Multi-layer visualization

Figure 3: Multi-matrix visualization

In the multi-matrix display, changes are reflected primarily in terms of color. Blue and red show dependencies, corresponding to an approval-disapproval heat map: the upper, blue, triangle, is associated with approval, while the lower, red, represents failure. When the mouse moves over a triangle, related courses are depicted, both in terms of failure and success (Figure 3, depicts courses related to *am1* in the first semester). When a course is selected, detailed information on that course is displayed in the rectangular panel between the two views. When a pattern is selected from this list, the corresponding information is reflected in terms of visual connectors (multi-layer view) and matrix cells (multi-matrix view).

EduVis enables course comparison. Clicking on a course locks it selected, keeping its information from being cleared with mouse movements. Moving the cursor over another course shows its information, allowing simultaneous comparison. The matrix view provides filtering mechanisms. When a course has been previously selected and another, with common relations with the first, is selected, restricts the visualization to courses with which both selected courses share one or more patterns, as shown in Figure 1, making a restriction on the initial set of interrelations of the first selected course. It is thus possible to add restrictions indefinitely.

Conclusion and future work
We created EduVis, a visualization which combines different interaction mechanisms to display dependencies among courses of an academic program. Preliminary evaluation has suggested the creation of mechanisms that allow users to learn how to interact. Furthermore, the integration of the two different prototypes needs further adjustment, particularly regarding the use of color, to increase consistency, as well as the readjustment of size and placement of visual artifacts. Nevertheless, the aforementioned evaluation has shown that the system displays educational patterns that were not evident otherwise, showing information that is, in fact, relevant.

References
[1] Gao, Y. A study on mass higher education in the world-based on comparative perspectives. In *International Conference on Education and Management Technology (ICEMT)* (2010), 528–530.

[2] Mazza, R., and Botturi, L. Monitoring an online course with the gismo tool: A case study. *Journal of Interactive Learning Research 18*, 2 (April 2007), 251–265.

[3] Mazza, R., and Dimitrova, V. Generation of graphical representations of student tracking data in course management systems. In *Ninth International Conference on Information Visualisation* (2005), 253–258.

[4] Ware, C. *Information Visualization: Perception for Design*. Elsevier, 2012.

[5] Wortman, D., and Rheingans, P. Visualizing trends in student performance across computer science courses. In *Proc. 38th SIGCSE Technical Symposium on Computer Science Education* (2007), 430–434.

[6] Xiaohuan, W., Guodong, Y., Huan, W., and Wei, H. Visual exploration for time series data using multivariate analysis method. In *8th International Conference on Computer Science Education (ICCSE)* (2013), 1189–1193.