World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

A Formal Approach for Instructional Design Integrated with Data Visualization for Learning Analytics

Douglas A. Menezes, Isabel D. Nunes, Ulrich Schiel

Abstract—Most Virtual Learning Environments do not provide support mechanisms for the integrated planning, construction and follow-up of Instructional Design supported by Learning Analytic results. The present work aims to present an authoring tool that will be responsible for constructing the structure of an Instructional Design (ID), without the data being altered during the execution of the course. The visual interface aims to present the critical situations present in this ID, serving as a support tool for the course follow-up and possible improvements, which can be made during its execution or in the planning of a new edition of this course. The model for the ID is based on High-Level Petri Nets and the visualization forms are determined by the specific kind of the data generated by an e-course, a population of students generating sequentially dependent data.

Keywords—Educational data visualization, high-level petri nets, instructional design, learning analytics.

I. INTRODUCTION

ID is the step of the construction of online courses in a Virtual Learning Environment (VLE), where customization questions, pedagogical and instructional techniques are specified [1]. The support for the planning of this step based on past instructional experiences is known as Learning Analytics.

VLEs are responsible for providing content, means of interaction and activities to people who are geographically distant or with unavailability of common time. However, most VLEs do not provide support mechanisms for ID, and the planning, construction and follow-up processes are carried out separately [2]-[4].

High-level Petri nets can be used to meet this need. These nets have a mathematical and formal basis that allows the development and monitoring of systems [5]. More precisely, high-level Petri nets, such as Colored Petri Nets [6], allow the modeling of activities, the specification of path constraints, and the individualization of marks, creating an environment suitable to planning (Prior to the course) and monitoring activities performed by students (at run time) in VLEs.

The High Level Activities Net (HLAN) is a notation that brings up the necessary features to meet the requirements of ID planning and its follow-up at run time. The HLAN uses as a formal basis a special class of Colored Petri Nets [7],

D. A. T. de Menezes and U. Schiel are with the Federal University of Campina Grande, Campina Grande, PB, Brazil (e-mail: dougglas2002@gmail.com, uschiel@gmail.com).
I. D. Nunes is with the Federal University of Rio Grande do Norte, Natal, RN, Brazil (e-mail: bel@imd.ufrn.br).

supplying their needs and characteristics and thus becoming specific to the type of process to be modeled.

A HLAN allows you to specify activities and their interdependencies, and the status of individuals and groups realizing these activities. Such characteristics are fundamental in an ID process. Integrating such a net in a VLE, the model allows monitoring students' activities during a course execution, thus identifying students at risk and badly planned DI points. Tracking runtime activities also allows you to store details about ID execution over time, creating a history of the course. This history can be analyzed by Learning Analytics algorithms [8], producing information pertinent to the improvement of future ID.

Learning Analytics [9] is the approach that uses the data generated during the application of a course in order to understand and improve the learning process. A valuable support for this understanding and improvement may be the visualization of aggregates of the data as adequate graphics, allowing a quick interpretation and perception of important indicatives of the courses and their students. This can facilitate the analysis of the data, increasing the cognition and understanding of the generated information [10].

The Authoring Tool for Instructional Design (ATID) [11] intends to support the building of the ID of a new course, using the HLAN notation and features, and to go along its realization, keeping all data generated by the participants. An adequate use of this data can point out significant indicators about the performance of a class, such as the level of income and motivation of the students, factors that can determine directly the performance of the class [12].

Information obtained through Learning Analytics techniques should be viewed in an efficient and easily understood manner by teachers, students and coordinators. The visualization of educational data allows to identify important knowledge about the productivity of classes, the level of commitment and motivation of the students, and factors that can influence directly the learning process of the students. Based on this fact, the computational data generated during a course and the information that was processed can be used by specialists. In order to allow a fast interpretation and perception of important indicatives of the course and their students, the presentation of the information as a summary in the form of adequate graphical visualizations is mandatory.

This paper presents an authoring tool for the creation of High-Level Activity Nets, its integration in a VLE and connected with a Data Visualization Tool which follows the

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

running of a course and analyses the quality of past achievements of a course.

II. ACTIVITIES NETS AND DATA VISUALIZATION

This section describes the main concepts used in this work. First, the HLAN concept, with its symbology, is explained.

The concepts of Data Visualization for Learning Analytics are then presented, since it was used to analyze the information obtained through the course.

A. HLAN

A HLAN is a graphical model used to plan and follow the process of developing a project [13 (5-icalt)]. A HLAN can be used as a structuring tool for ID, as it contains the necessary elements to model activities (simple and compound), events (criteria for activating an activity), artifacts and allows to describe parallel and concurrent flows of a course.

An HLAN is a high-level Petri net with a more expressive notation, since it contains elements for the representation of composite activities, events, and artifact repositories, in addition to that, each activity can have time attributes [13].

An Activities Net can be modeled with the graphical elements shown in Table I. The meaning of the links of activities and transitions is as follows:

- A link from a basic activity to a transition means that its execution is a necessary condition for enabling the transition. It will be enabled to fire when all incoming activities are marked on the outer ring. Transition triggering removes the hallmark of input activities.
- A link from a composite activity to a transition means the same as for the basic activities, except for a special transition called the sub-net input transition (snit). The snit has a composite activity as input and is enabled when it is marked on the inner ring. The execution of snit

moves the mark to the middle ring of the composite activity.

- A link from a transition to an activity means that the transition trigger marks the activity on the inner ring. An exception is a special transition to a composite activity, called the sub-net output transition (snot). Its firing moves the composite activity mark from the intermediate ring to the outer ring.

An example of a HLAN is shown in Fig. 1. The HLAN shows the sequencing of activities, including the repositories required for its execution. In the example, after the third activity the student has two alternative ways to go further in the course, one of them has two independent activities to be realized.

TABLE I
BASIC ACTIVITIES NET GRAPHIC ELEMENTS

| Symbol	Quantity
	Basic Activity - A mark on the inner circle indicates that the activity is being performed; A mark on the outer circle indicates an activity finished, enabling transitions to other activities that depend on it. An unbranded element indicates that the activity is disabled.
	Composite Activity - is an activity performed by a subnet associated with it. A mark on the inner circle indicates that the activity can be started. A mark on the middle circle indicates that the subnet is running, and a mark on the outer circle indicates that the subnet has terminated its activities.
	Repository - stores artifacts produced by activities and that can be used by later activities.
	Transition - is the transition between finished activities (or subnets) and subsequent activities.
	Establishes links between activities, events, artifacts, and transitions.
	Represents a markup on the network. It determines the state of the activity (running, disabled, or terminated).
	Event - Allows you to model the conditions necessary for triggering a transition.
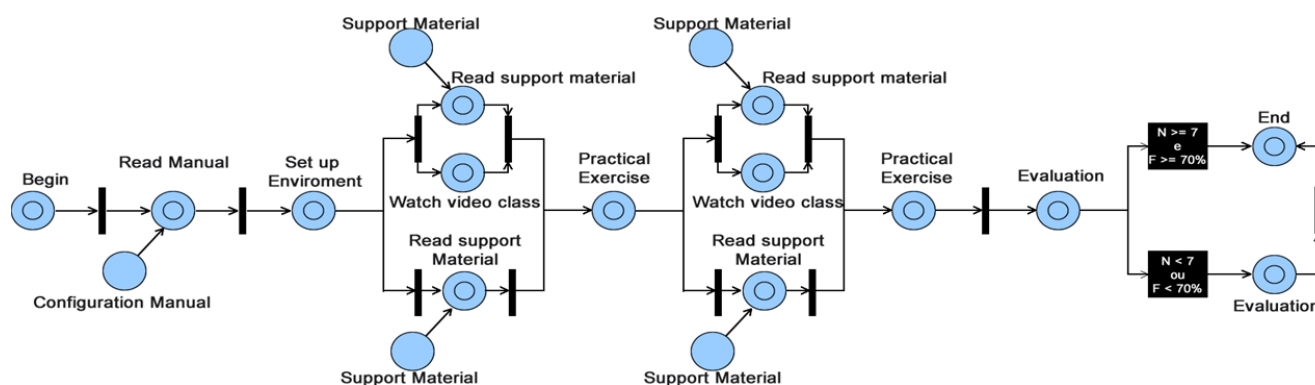
Fig. 1 Example of a HLAN

An ID is well formed when the corresponding HLAN that represents it does not contain certain connections that may compromise its execution. The situations in which an Activities Net is compromising are: when it is not possible to continue network execution (deadlock) and when an activity becomes useless and does not have to be terminated.

Since a HLAN can be mapped to a Coloured Petri Net the corresponding Reachability Graph of this net allows us to check if the HLAN is alive, that is, the network comes to an end (no deadlock occurs) and it does not leave unfinished activities. The Reachability Graph of possible states of the net must be a lattice, allowing a single final state, representing the end of the course [14]. This check is necessary so that IDs are well planned and do not cause problems during the running of a course.

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

B. Data Visualization for Learning Analytics

The planning of a course is the purpose of the ID process [15]. In the activity of analysis one must know the students, their context and the activities that best fit their profile. From this, the organization of the course, and how it will be managed, must be defined. Finally, a formative evaluation takes place and the revision of the instructions is carried out. Learning Analytics techniques allow, in addition to analyzing information about student performance, extracting and analyzing information about the progress of a discipline [9].

From the analysis performed on the student profile data, the defined ID model, and the learning information produced during the course, the teacher can introduce improvements in his discipline. Also, historical data from past editions of the course can support the improvement of ID.

Data visualization, in essence, consists of a visual representation of a set of data in the form of an image within a mental model, which corresponds to a mechanism of thought by which man tries to understand how objects are in the real world [16].

As in textual representations, projected images are designed to create a kind of visual correlation with a situation expressed by a mass of data. Data visualization also has advantages in the degree of human perception and can significantly improve both the quality and the quantity of information that is presented through the generated computational images [16]-[18].

In an educational context, data visualization should reflect the sequential nature of the generated data, each of them reporting quality aspects of a learner. Therefore, visualization will have a structure of a directed coloured graph, specifically a tree.

III. Authoring Tool for Instructional Design

The use of a HLAN integrated in a VLE permits to follow the activities of the students during the execution of a course, allowing to identify students at risk and critical points of the ID.

A tool for this e-learning process, called ATID - Authoring Tool for Instructional Design [11] was created, which allows the creation of HLANs integrated into a VLE that provides student information during the course.

The ATID allows to create the structure of the ID of a certain course, defining all possible sequencing of course activities, priorities, validation rules for each transition and execution periods. The created net can be used during the course execution storing all generated data for on-line assessments or as the basis for a new course model. Validation rules for allowed steps forward in the course are defined as Guard Conditions assigned to the transitions of the Activities Net.

Fig. 2 shows the layout of the tool, in its first version, with the editing of a network of activities.

IV. Case Study

This section presents a fictitious course that contains alternate and parallel activities. The activities are defined based on activities carried out by actual distance courses.

The HLAN of that represents the course as a whole can be seen in Fig. 3. In this structure, eight alternative paths were defined for the course. The purpose of this course is to verify how the different profiles of students behave different in the same course.

Students have eight different possibilities for completing the course, including alternative paths and also the need to perform the final evaluation again if they have not reached the required grade. Thus, each combination of these alternatives is a different way of doing the course. The analysis of each path separately allows a better identification of the possibilities that the students possess.

The completion of the case study depends on course performance data. Since such information is difficult to access, in this work, a learning simulation was performed.
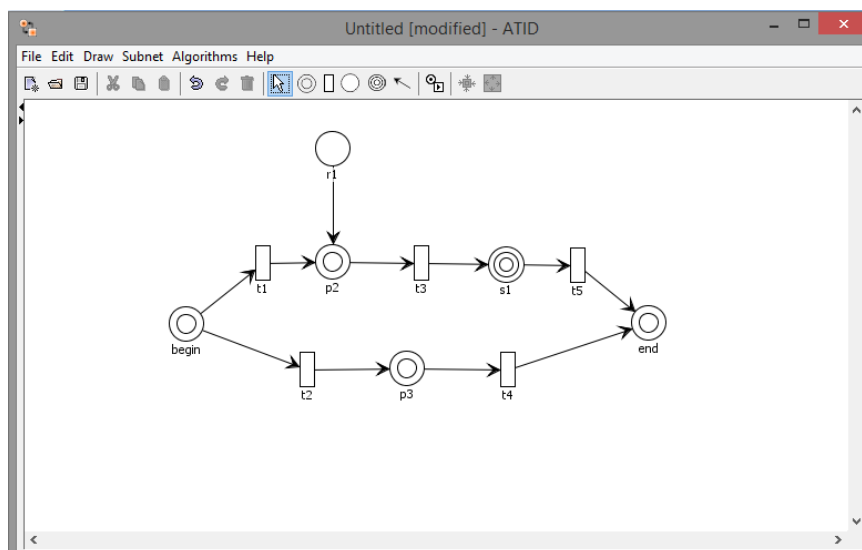

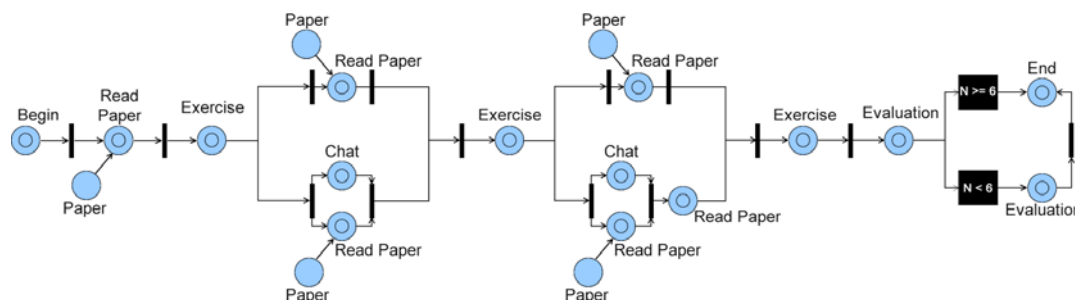Fig. 2 Main Window of Atid for the Creation of Activities Nets

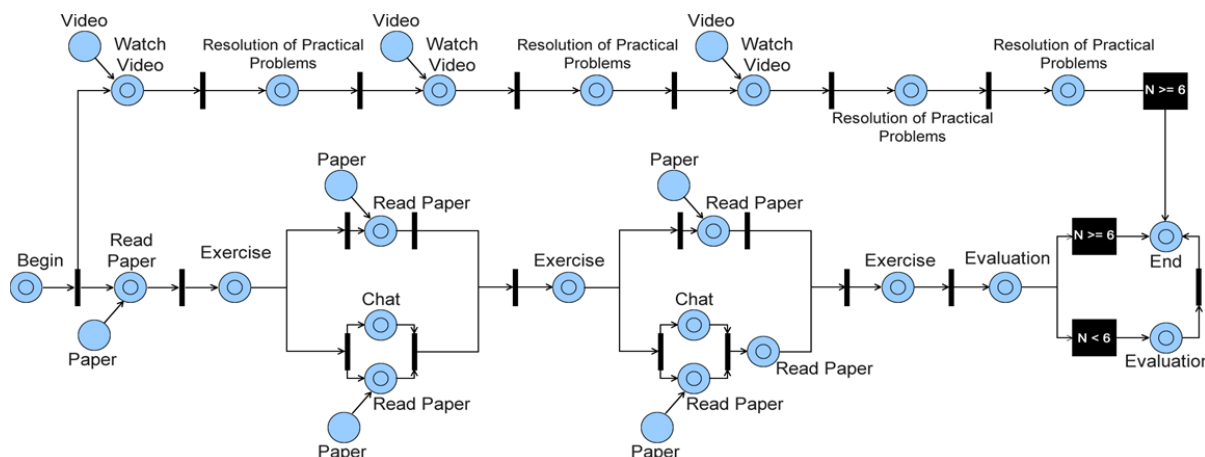World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

Fig. 3 Course Activities Net

Fig. 4 Improved HLAN

The Student Profile is defined from the combination of the four dimensions defined in the Learning Style [19]. In total, 16 profiles were defined (combination of two possibilities for each of the four profiles), which are: P1 (active, sensitive, visual and sequential), P2 (reflexive, visual, and sequential), P4 (active, sensitive, verbal and sequential), P5 (active, sensitive, visual and global), P6 (reflexive, intuitive, visual and sequential), P7 (active, sensitive, verbal and global), P9 (reflexive, sensitive, verbal and sequential), P10 (reflexive, sensitive, visual and global), P11 P15 (reflective, intuitive, verbal and global), P14 (reflective, intuitive, verbal and global), P13 (reflexive, intuitive, verbal and global).

The evaluation of the Case Study was performed through 20 repetitions for each of the 16 profiles of students defined for each Path proposed in this course.

Each activity receives a number for its identification in the analysis. The eight paths are identified as C1 to C8.

Profiles P1, P2, P5 and P10 have been disapproved in all paths, indicating that even in the course providing alternative paths there is still a need to reach groups of students who have difficulty in taking the course because they have a learning style different from the style adopted by the course.

Profiles P7, P8, P9, P12, P13, and P16 were approved in all paths. Such profiles are strongly connected to the profile of the course.

A. Improvements in the ID

From the performed evaluation, the more adequate path id the path C1, in which 12 of the 16 profiles were approved.

Particular attention should be given to the profiles P1, P2, P5 and P10 that did not fit in any path provided by the course. For such profiles, it is necessary to propose a course focused on their characteristics, which have in common being sensitive and visual, which is not attended to by the activities proposed in the eight paths of the course.

In order to meet the profiles that did not get approval in the course, the analysis performed in the projected paths, one has:

- Activities c0 and c1 were those that had the highest number of disapprovals and should be altered (as shown in Fig. 3);
- Path C1 was the one that obtained the highest number of approvals, being considered for the improvement of the course.

Fig. 4 shows the ID proposal for the Fictitious Course considering the analyses performed with Learning Analytics and the results obtained.

The ID presented in Fig. 4 contains the insertion of a new path for approval of student with profiles P1, P2, P5 and P10.

V. THE DATA VISUALIZATION MODEL DaVID

The data visualization model DaVID-Data Visualization for the ID was developed with the aim of presenting a clean, intuitive and easy-to-use interface that allows the analysis of a course both during its running.

An ID offers several ways to carry out a course. These paths are presented as an acyclic-directed graph with a single ancestor (tree), whose nodes represent activities and edges the

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

sequencing relationships between activities (Fig. 5).

The color assigned to each node represents the performance of the group of students in that activity, that is, the color tone associated with a note is taken for each student and the average is calculated.

For students who drop out the course and for activities that do not have concepts (note), the assigned color will be white (Fig. 6). At the end of the last activity a node is added for each student who completed the course through the corresponding path (Fig. 6). This is an optional function, since if a class has too many students enrolled, this visualization would not be adequate, and a single note represents all students.
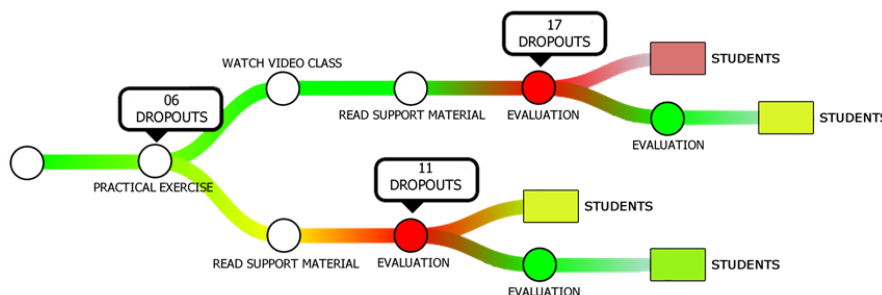

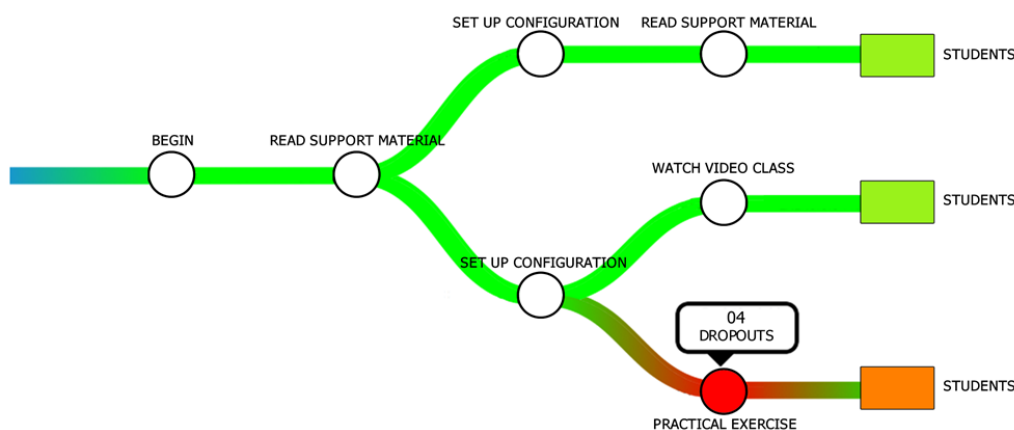Fig. 5 Partial View of the Possible Paths of ID

Fig. 6 Identification of the Dropping Student

Thinking about very large student groups, the module can be configured so that it does not display all students who finished the course this way but only one node representing all students who have completed the course along this path (Fig. 7). The color of this node would be in accordance with the criteria presented previously for the performance representation of the group of students.

Clicking on the node representing a group of students a screen will be displayed listing all the students who have completed the course by this path. It can be seen in Fig. 8 that for each of the students who completed the course on Path 1 a graphical representation is shown indicating the path traveled, where each circle represents an activity and the color represents performance.

If the teacher wants to see more details about the performance of a specific student, just click on the student's name that will be displayed detailed information about each of the activities performed, including notes, time spent on each activity, and personal information such as: Age, sex, school of origin, etc. (Fig. 9). In the screen that presents the details of

the student, two tabs are available that aim to detail the performance of the student in each activity. In the first tab, Bar Graph, bars are provided that represent student performance in each activity.
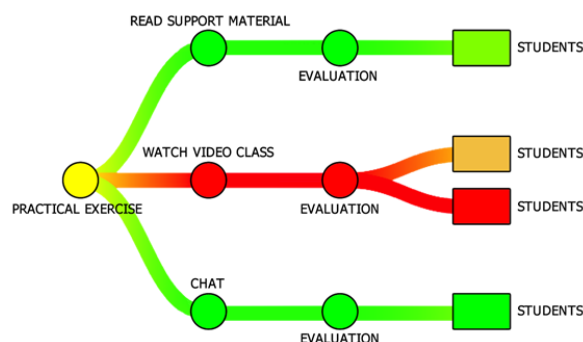

Fig. 7 Aggregated Student Information

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

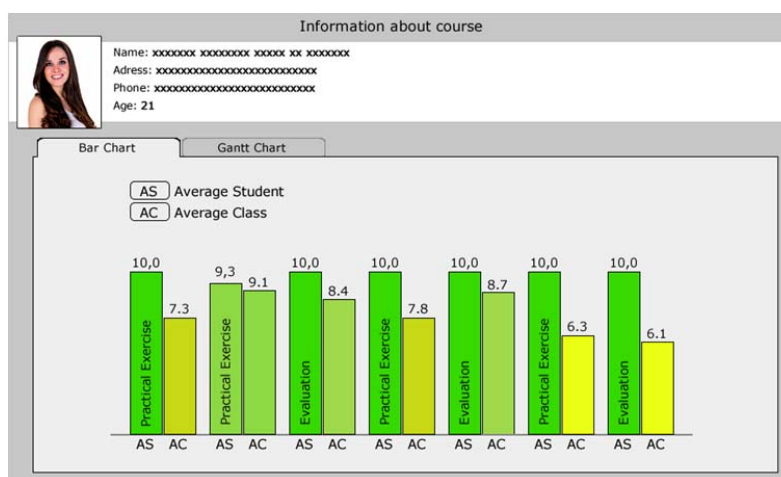Fig. 8 List of Students Who Completed the Course by a Specific Path

Fig. 9 Student Details – Bar Chart

In Fig. 9, two bars are presented side by side for each activity. The first represents the concept of the student in an activity and the second is the average obtained by the group in the same activity.

In the second tab, a Gantt Chart gives details about the completion times of each activity. This graph shows the time spent in each activity through a timeline, in addition to showing details related to the execution of each activity, such as the start date of the activity, when it was completed, the number of accesses, the number of Days for the activity and note (if this activity has a note) (Fig. 10).

In addition to listing students by specific paths, it is also possible to list all enrolled students who have completed the course, regardless of path. The data visualization module allows the teacher, through the complete listing, to select a group of students according to a pre-established criterion, in order to cross the information obtained by this group of students to find specific patterns. The teacher can order the list of students by: better performance, worse performance, longer

time to complete the course, shorter time to complete the course.

It is also possible to use a filter, where the teacher can select the students by range of notes. The order functionalities and filter by note interval are found on the left side of the screen, as shown in Fig. 8.

Taking as example the possibility of the teacher ordering the list of students by the final average, he can select a specific group of students, For instance, selecting only the first six students, it is possible to generate a new visualization with the paths of these students (Fig. 11). The purpose of this functionality is to identify the success paths from the students who obtained the highest income. The same thing can be done to identify the paths that present problems, the paths of success and failure by profile, age, sex, etc.

The graphical representation of an ID is given by the corresponding HLAN of the course in question. The Activities Net shown in Fig. 1 represents the ID of a specific module in a computer programming course. This HLAN shows that the ID

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

allows the student 16 ways to complete this module, if the ID takes into account the order of parallel activities in each of the bifurcations.

To interact with the data visualization module, it is necessary to click on a certain node (activity), where it will show information about the number of students that went along the path to this activity, as well as the number of students that got high score (>= 7), as well as those with a medium grade (5 <= and < 7), a low scale and those who dropped out (Fig. 12).
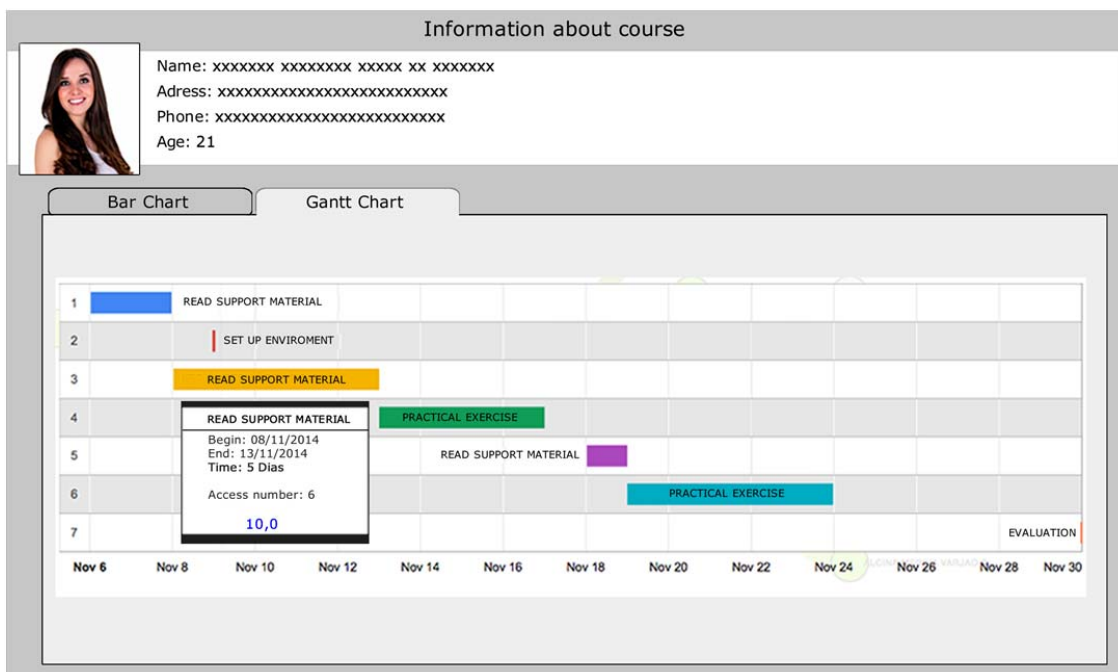
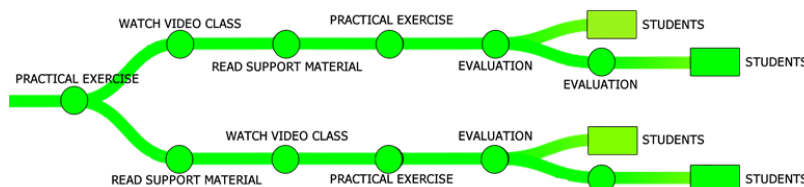

Fig. 10 Student Details – Gantt Chart

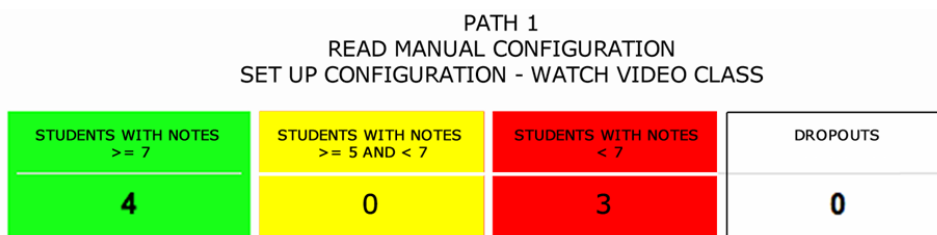Fig. 11 Successful Paths Based on Previous Selection

Fig. 12 Students Achieving in an Activity

The second display option corresponds to a line chart, which shows the performance of the student group up to a specific activity (Fig. 13). Each color represents a student. If the teacher wants to see the specific performance of a student, just click on the name that corresponds to this student and will be evidenced the performance of this student. This type of visualization has a relevance when you want to observe a trend of notes per student group by a specific path or set of activities.

The data visualization module has a customizable area that allows users to filter the data, taking into account the student profile, generating a new view based on this filter. To do this, simply access the settings menu and enter the necessary parameters to generate the visualization according to the chosen profile (Fig. 14). This selection is important to identify trends in choosing ID because you can compare success paths and problematic paths based on specific profiles.
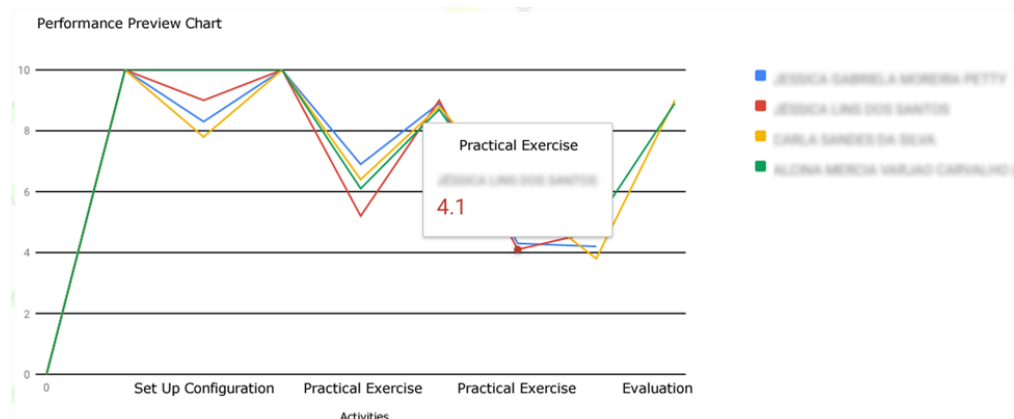
World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

Fig. 13 Line Display of Individual Students

As stated earlier, the last nodes represent students who have reached the end of the course, either individually or as a group. Therefore, it is also possible to obtain information about a student or group of students. The process is the same presented for the activity nodes. Clicking on the node that represents the student will display the Gantt chart, which shows the time spent per student in each activity, as shown in Fig. 10.

The DaVID model can also be used during the running of a course. In this case, a partial graph of the situation of the students can be visualized at a certain moment (Fig. 15) and the time-line until a certain time t.

Clicking on the node that represents the student group of a specific path, a timeline chart is displayed, which shows, chronologically, how the students are doing activities on that path (Fig. 16). This type of visualization shows the progress of the activities per student, highlighting those that are in the term, delayed or advanced.

Because the data visualization module evidences information through graphics and color, the use of color may be a problem for users with color blindness. To solve this problem, the module allows the use of a suggestive emoji to identify the success or otherwise of that activity (Fig. 17).
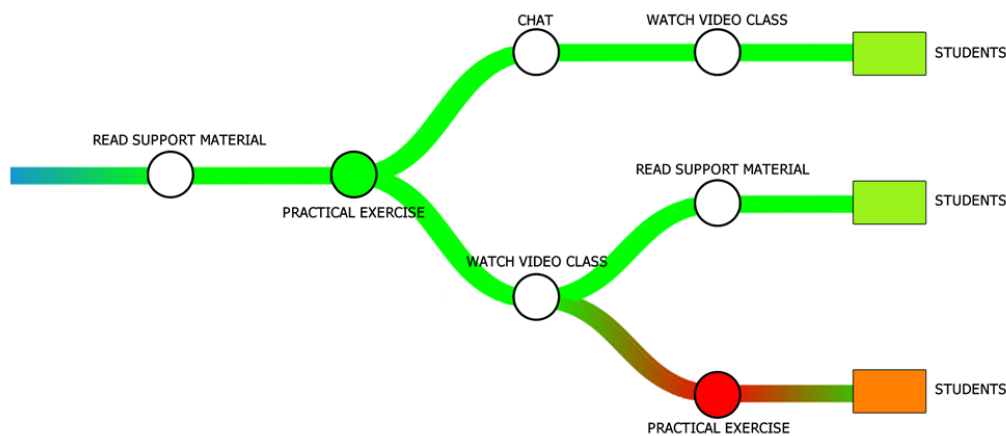

Fig. 14 Settings Menu

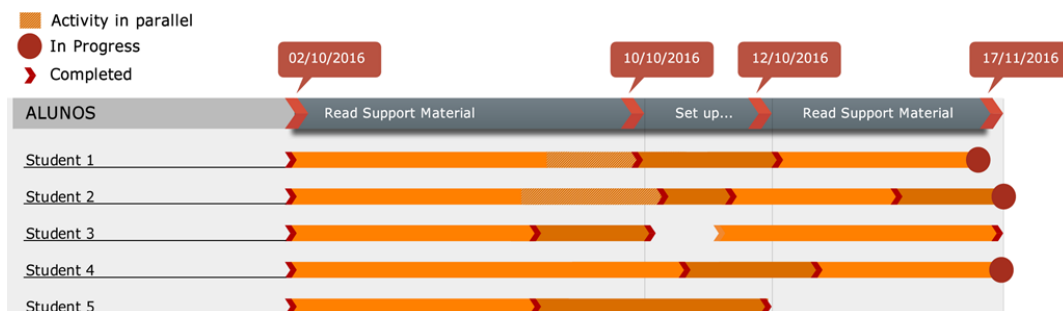Fig. 15 Partial Graph of Course Progress

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
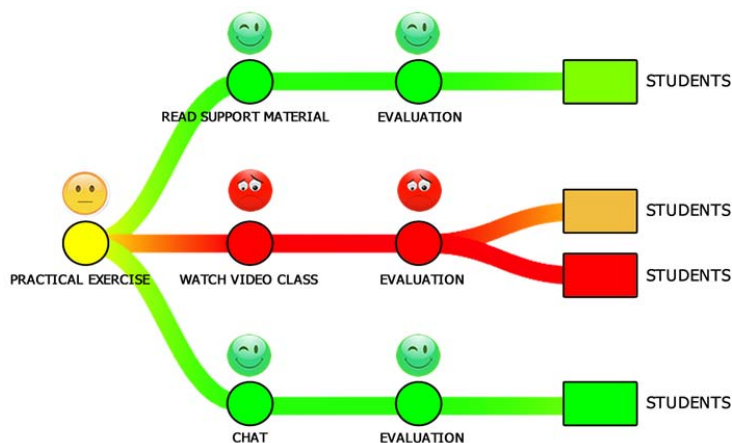Vol:11, No:10, 2017

Fig. 16 Time Line Chart

Fig. 17 Data Visualization with Support for Color Blind

VI. CONCLUSION

This paper has presented a proposal for the integration of tools for the construction of ID structures with the use of Learning Analytics techniques for the evaluation and improvement of these structures. The ID structure allows a precise monitoring of course applications and the corresponding authoring tool keeps all generated data in a database. Since the volume of this data may be large, a visualization tool has been proposed to synthesize the information contained in the data and present it visually.

Since the HLANs, as model for the ID, are based on High-Level Petri Nets, a completely sound specification of a course structure is guaranteed. The existence of a tool for the creation and using of HLANs, allows to keep all data generated by a course in a database. The visualization tool follows the sequential order of a course execution giving a precise mirror of the quality of the course, allowing route corrections and planning of improved editions of the course.

The visualization tool has been validated successfully alongside the application of a computer programming course.

REFERENCES

[1] B. Fetaji and M. Fetaji "E-Learning Indicators: A Multidimensional Model for Planning Developing and Evaluating E-Learning Software Solutions" In: E-Learning, experiences and future. Edited by Safeeullah Soomro and published by In-Teh. Pag 1- 34. 2010.
[2] A. Vahldick, R. Santiago and A. L. A. Raabe "Aplicação das Técnicas de Projeto Instrucional 4C/ID na Produção de Objetos de Aprendizagem em Conformidade com o SCORM usando um Software Livre como Ferramenta de Autoria". Revista RENOTE – Novas Tecnologias na Educação. V. 5, nº 2. 2007.
[3] A. Filatro, R. Boaretto and C. Nunes. "Learning Activities Editor Tool." TDIA – III Workshop. Pág. 152-154. 2006.
[4] T. Vantroys and Y. Peter "COW, a Flexible Plataform for the Enactment of Learning Scenarios." Lecture Notes in Computer Science. 2003.
[5] S. I. D. de Pádua, A. R. Y. da Silva, R. Y. Inamasu and A. J. V. Porto. "Redes de Petri aplicadas aos Sistemas de Gerenciamento de Workflow." XXII Encontro Nacional de Engenharia de Produção. Curitiba – PR. 2002.
[6] K. Jensen. "Coloured Petri Nets – Basic Concepts, Analysis Methods and Practical Use." Volume 1. Second Edition. Springer-Verlag Berlin Heidelberg New York. 1997.
[7] I. D. Nunes and U. Schiel. "Design Instrucional e seu acompanhamento em tempo real utilizando Rede de Atividades". In: Simpósio Brasileiro de Informática na Educação, 2011, Aracajú. Simpósio Brasileiro de Informática na Educação, 2011.
[8] T. Elias. "Learning Analytics: Definitions, Processes and Potential." Available in: http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPoten tial.pdf. 2011. Accessed 15 January 2017.
[9] R. Ferguson, "The State of Learning Analytics: A Review and Future Challenges." Technical Report KMI-12-01, Knowledge Media Institute, The Open University, UK, 2012, http://kmi.open.ac.uk/publications/techreport/kmi-12-01. Accessed 16 January 2017.
[10] M. Scaife and Y. Rogers. "External Cognition: how do graphical representations work?" International Journal of Human-Computer Studies, 45(2): 185-213, 1996.
[11] C. Medeiros, W. Balbino, D. Menezes, A. Barbosa, I. D. Nunes and U. Schiel, "ATID - Authoring Tool for Instructional Design" In Anais dos Workshops do Congresso Brasileiro de informática na Educação (Vol. 4, No. 1, p. 383), 2015.
[12] A. G. G. Cavalcanti, N. N. dos Santos, J. L. Cavalcanti and A. S. Gomes, "Mineração e Visualização de Dados Educacionais: identificação de fatores que afetam a motivação de alunos na educação a distância" XIV Escola Regional de Computação Bahia, Alagoas e Sergipe (ERBASE), 2014.

World Academy of Science, Engineering and Technology
International Journal of Educational and Pedagogical Sciences
Vol:11, No:10, 2017

[13] I. D. Nunes and U. Schiel, "Using High Level Activities Net for Learning Analytics of Instructional Design" Proc. IEEE International Conference on Advanced Learning Technologies (ICALT 2014), July 2014, doi: 10.1109/ICALT.2014.114.

[14] I. D. Nunes and U. Schiel. "Adaptive Instructional Design using High Level Activities Net" E-Learn - World Conference on E-Learning in Corporate, Government, Healthcare & Higher Education, 2013, Las Vegas. World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (ELEARN) 2013, 2013.

[15] A. Filatro, "Design Instrucional na Prática". Pearson Education do Brasil, 2008, 173 páginas.

[16] C. Ware. "Information Visualization: perception of design." Elsevier, 2012.

[17] S. K. Card, J. D. Mackinlay and B. Shneiderman. "Readings in Information visualization: using vision to think." Morgan Kaufmann, 1999.

[18] C. Chen. "Information visualization: Beyond the horizon. 2nd Edition" Springer, 2010.

[19] R. M. Felder. and L. K. Silverman. "Learning and Teaching Styles". In: Engineering Education. Engr. Education, 78(7), 674-681. 1988.