A computer approach to mathematics curriculum developments debugging

Angélica Martínez-Zarzuelo, Eugenio Roanes-Lozano, Mª José Fernández-Díaz
Universidad Complutense de Madrid, SPAIN

Sequencing contents is of great importance for instructional design within the teaching planning processes. We believe that it is key for a meaningful learning. Therefore, we propose to formally establish a partial order relation among the contents. We have chosen the binary relation “to be a prerequisite” for that purpose. We have applied this approach to the mathematical contents of the compulsory Secondary Education of the Spanish educational system and the information obtained has been modeled as a graph. The amount of contents considered (814) and the number of ordered pairs in the order relation considered (17,782) has produced a big graph. In order to work effectively with that amount of data, we have used software specialized in network analysis. More precisely, we have used the software packages Pajek and Gephi. This software, together with the use of techniques borrowed from graph theory, has allowed providing a tool for debugging curriculum developments (similar to rule based expert systems verification).

Keywords: curriculum development debugging; graph theory; instructional design; network analysis; organization of mathematical contents; sequencing contents.

INTRODUCTION

This work proposes and validates a methodology based on the application of techniques borrowed from graph theory (Kaufmann, 1976; Wilson, 1983; Abellanas & Lodares, 1990; West, 2001; Aldous & Wilson, 2003; Biggs, 2003; Chartrand & Zhang, 2005; Balakrishnan & Ranganathan, 2012) and specialized software in network analysis to debug proposals of organization of educational contents in teaching planning.

By debugging we mean "verifying that the educational contents sequencing proposed is correct" (something similar to what is done in rule based expert systems verification). This way we believe we contribute to the development of models of instructional design. In order to prove the viability of this methodology, it has been applied to the mathematical knowledge (curriculum developments) of the education stage “E.S.O.” (compulsory Secondary Education) of the Spanish

Correspondence: Angélica Martínez-Zarzuelo, Universidad Complutense de Madrid, Rector Royo Villanueva, s/n 28040 Madrid, Spain E-mail: angelica.martinez@ucm.es

Copyright © 2016 by the authors; licensee iSER, Ankara, TURKEY. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original paper is accurately cited.

ISSN: 1305-8223 http://iserjournals.com/journals/eurasia
educational system. However, the nature of this methodology makes it possible to easily extend it to other areas of knowledge, as well as to other educational stages. This work summarizes part of the Ph.D. Thesis of the first author, whose advisors were the other two authors (Martínez-Zarzuelo, 2015).

LEGAL FRAME

We shall follow somehow a down to top design: we shall begin by analyzing the example selected (Spanish compulsory Secondary Education) in order to have a panoramic view of the object of study that allows developing the theoretical/computational approach.

The present Spanish K-12 educational system distinguishes 4 educational stages:
- Kindergarten,
- Primary Education,
- compulsory Secondary Education ("E.S.O."), for 12- to 16-year-olds,
- non-compulsory Secondary Education ("Bachillerato").

This study focuses on the third stage. The reasons for choosing this educational stage have been, firstly, to be a compulsory educational stage. This allows performing the study on a stage that involves a huge amount of students of the Spanish educational system. Secondly, it is the last compulsory stage. This fact makes contents more complex and diverse. Consequently, it makes the data of the study richer.

We have analyzed the details of the mathematical subjects of the last three laws regulating compulsory Secondary Education in Spain:
- LOGSE (Ley Orgánica de Ordenación General del Sistema Educativo, 1990; Real Decreto 1007/1991, 1991),
- LOE (Ley Orgánica de Educación, 2006; Real Decreto 1631/2006, 2006), and

These core curricula detail the objectives to be reached, the contents to be taught (not in detail), and the competencies to be developed. It also makes some suggestions on the methodology and assessment (what does not affect this study).

THEORETICAL FRAME AND RELATED WORKS

A correct core curriculum development is generally considered to be the key to a more practical and proactive regulatory framework (Schubring, 1987; Leuders, Barzel & Hußmann, 2005; Goñi, 2011). This characteristic enables an approximation between teaching theory and practice. Although the educational laws usually leave a good degree of freedom to curriculum developers (of textbooks, educative projects, etc.), they usually just translate the suggested ways of organizing suggested by the laws. Moreover, the format of textbooks leads to a contents sequencing mainly
linear, in accordance with their appearance in the textbooks (Ausubel & Barberán, 2002).

Nevertheless, there are many other possible ways to sequence contents. For instance, Zabalza (2009) proposes different kinds of sequencing based on criteria such as the importance of the contents involved in a sequence and the time required for teaching them.

From our point of view, the most interesting criterion is the one based on meaningful learning (Ausubel, 1963; Ausubel, Novak & Hanesian, 1976; Moreira, 2000), where to be an anterior or ulterior content is taken into account.

Regarding this issue, Rodríguez (2008) asserts that “no se puede desarrollar aprendizaje significativo en el alumnado con una organización del contenido lineal y simplista” (“students cannot develop a meaningful learning if the content is organized in a linear and simplistic way”) (p. 31). In this sense, two different sequencing for teaching could be considered, for example, the contents Subtraction, Multiplication and Division from the content Summation (Figure 1).

![Figure 1. Two alternative contents sequencing (Rodríguez, 1983, p.68).](image)

On the other hand, the internal logic structure of the subject to be taught and a balanced equilibrium between procedural skills and conceptual understanding (Drijvers, Goddijn & Kindt, 2011) are key to plan its contents sequencing. Mathematics has a very special internal logic structure: contents are based one on another. Although this is possibly the reason for the high percentage of students that hate this subject, it is also the key for organizing and sequencing contents, and, consequently, for a meaningful learning. That is why this subject was chosen to exemplify and validate the methodology proposed in this work.

Nevertheless, it is clear that the idea developed here could be applied to other subjects. In this context we can find Ausubel’s meaningful learning theory (Ausubel, 1963; Ausubel, Novak & Hanesian, 1976; Moreira, 2000; Ausubel & Barberán, 2002), as this theory holds that new information is meaningful as far as it can be related with information already known.

Regarding the internal logic structure of mathematics as a subject, there is a wide consensus about the importance of understanding it as a network of interrelated concepts and procedures, and not as a collection of rules and isolated facts (Brinkmann, 1999, 2001, 2003, 2005).

The technique underlying conceptual maps (Novak & Gowin, 1988) is well suited for graphically representing knowledge. The reason for this is that it can be used for representing, experimenting, perceiving and learning the network structure of
mathematics (Brinkmann, 1999, 2003; Mwakapenda & Adler, 2003; Brinkmann, 2005). In addition, authors such as Sherborne (2014) discuss how the use of conceptual maps could help curriculum developers and teachers.

Brinkmann (2001) proposes modeling the mathematical connections using a graph too. Nevertheless, how to develop this process is not fully detailed. She underlines the necessity to avoid considering a large number of mathematical objects for the sake of clarity. Both aspects have been addressed in the methodology developed here.

This methodology also inspires tridimensional expert maps (Pérez, 1998; Suero, Montanero & Montanero, 1999; Pérez, Suero, Montanero & Montanero, 2000a, 2000b), that are considered an extension of conceptual maps.

Conceptual maps and tridimensional expert maps are the starting points of the structure used in this work.

Reigeluth’s elaboration theory (Reigeluth, 1979; Reigeluth & Stein, 1983; Coll, 1987; Reigeluth, 1987; Reigeluth & Curtis, 1987; Reigeluth & Stein, 1987; Sacristán & Gómez, 1989; Coll & Rochera, 1990; Reigeluth, 1999; Reigeluth, 2013), itself based on conceptual maps and tridimensional expert maps, is also an important theoretical foundation for this study.

One of the purposes of this work is also sequencing and organizing educational contents. The underlying theory includes other theories in the field, such as Ausubel’s meaningful learning theory (Ausubel, 1963; Ausubel, Novak & Hanesian, 1976; Moreira, 2000), Gagnè’s learning theory (Gagnè, 1970, 1973) and Bruner’s spiral theory of learning (Bruner, 1963, 1966, 1991, 1998, 2003, 2004), thereby including its different perspectives about instructional designs (Bruner, 1966; Gagnè, Wager, Golas, Keller & Russell, 2005; Reigeluth, 2013).

Another technique for knowledge graphical representation, the pathfinder associative networks (Schvaneveldt, 1990), is also an inspiration. A recent line of research has applied “pathfinder associative networks” to the study of cognitive structures and has led to the development of the theory of nuclear concepts (Casas, 2002; Casas & Luengo, 2004a, 2004b). The key concepts of the latter theory are the “geographic” organization of knowledge, nuclear concepts, and minimum cost paths.

Finally, regarding the kind of structure used in this work (the graph), we could underline that it has been used by other authors for other didactic goals, such as: determining teaching itineraries (Alarcón, Fernández, González & Martínez, 2007), treating the academic curriculum as a complex system with nodes representing courses and links between nodes representing the course prerequisites (Aldrich, 2015), or suggesting a model for the solution of motion problems with an intelligent tutoring system (Nabiyev, Çakiroğlu, Karal, Erümit & Çebi, 2016).

DESIGN OF THE APPLIED STUDY

The key idea of our work is to firstly manually detect the contents involved in the curriculum proposed. This search is performed on the educational laws and the corresponding curriculum developments (textbooks). Secondly, the binary relation “to be a prerequisite” is manually established in the set of contents (once this set is determined).

The educative contents considered and the binary relation established can then be modeled using a graph structure. Graph theory has been applied in very different environments and diverse goals (Martín & Méndez, 2004; Leydesdorff, 2007; Puchades, Mula & Rodríguez, 2008; Cardozo, Gómez & Parras, 2009; Carlos, Gallardo & Colomer, 2011; Pino, Jiménez, Ruiz & Bailón, 2011; Gómez, Zarrazola, Montero & Yañez, 2012; Sánchez, Moratala & Sanz, 2012; Nieto, 2013; Wilson, 2013; Clough, Gollings, Loach & Evans, 2014; Romo, Vélez, Solís, Luna & Espinoza, 2015).
Nevertheless, as far as we know, it has not been previously applied to the goal of this work.

As the binary relation is established in a somehow subjective way, we consider the transitive closure of the given relation (the smallest transitive relation that contains the given one). This way the binary relation is completed (although subjected to errata in the data introduction). Finally, if we compute the transitive reduction of the latter graph, a canonical binary relation is obtained.

**ABOUT THE SOFTWARE USED**

We initially tried to “organize” the contents using pen and paper. The approach looked impractical soon after we began (Figure 2). The information was difficult to manage and it was almost impossible to introduce changes or alterations in it (like new contents, changes in the contents, changes in the binary relation). The first steps of the process were detailed in (Martínez-Zarzuelo, Roanes-Lozano & Fernández-Díaz, 2013).

![Figure 2. A pen and paper first approach.](image-url)
It was therefore clear that a network analysis piece of software would be a great help. After evaluating different ones, Pajek\textsuperscript{1} (Batagelj & Mrvar, 2014) and Gephi\textsuperscript{2} (Cherven, 2013, 2015) were chosen because of their power, their facilities to introduce data from Excel, and the possibility to exchange data between the two of them. Moreover, Pajek is free for noncommercial use and Gephi is open source, they do not require programming skills and can deal with graphs of high order and a big size.

Software specialized in visualization and analysis of information structured as a network in general, and these software packages in particular, have been widely used in a great variety of issues (White, Batagelj & Mrvar, 1999; Peñaranda, Quiñones & Osca, 2009; Camino, 2012; Conway & White, 2012; Raper, 2012; Álvarez, Kuz & Falco, 2013; Heymann & Le Grand, 2013). Nevertheless, they have not been used for the goal of this work.

**DEVELOPMENT OF THE APPLIED STUDY**

The binary relations "to be a prerequisite" and "to be an immediate prerequisite" have been defined in a theoretical way. Specifically, we have used the following definitions:

**Definition 1:** Let $\mathcal{C}$ be the set of mathematical contents of Spanish compulsory Secondary Education.

**Definition 2:** Let $c_1, c_2 \in \mathcal{C}$. $c_1$ is a prerequisite of $c_2$, denoted $c_1 \triangleright c_2$ if and only if understanding content $c_1$ is required to understand content $c_2$.

**Definition 3:** Let $c_1, c_2 \in \mathcal{C}$. $c_1$ is an immediate prerequisite of $c_2$, denoted $c_1 \triangleright\triangleright c_2$ if and only if understanding content $c_1$ is required to understand content $c_2$ and $\nexists c_3 \in \mathcal{C}$ such that $c_1 \triangleright c_3 \triangleright c_2$.

Once the two binary relations have been defined, the ordered pairs of mathematical contents (within the educational stage of Spanish compulsory Secondary Education) that define them are determined. In order to obtain these ordered pairs, we have used the guidelines given by the educational laws mentioned above and we have explored in detail 40 textbooks of the main Spanish publishers. More precisely, we have used ten textbooks for each of the four courses of the educational stage considered.

The definitions of each of the mathematical contents have been carefully analyzed. For example, we can find in one textbook that: "The set of reals numbers is formed by all the rational numbers and all the irrational numbers." Therefore: real number, rational number, irrational number $\in \mathcal{C}$ and:

- rational number $\triangleright$ real number
- irrational number $\triangleright$ real number.

The pairs of contents established this way are modeled using a digraph. When applied to the four courses of the compulsory Secondary Education in Spain, the digraph associated to the binary relation $\triangleright$, denoted $G = (\mathcal{C}, A)$, turns out to have 814 nodes and 3,925 arcs.

This digraph, as is initially introduced, is not really a graph, but a multigraph. It has multiple arcs because a kind of Bruner's spiral curriculum (Bruner, 1963, 1966, 1991, 1998, 2003, 2004) is adopted by the Spanish law and some arcs appear in different academic years. Therefore, a first step in the preprocessing is the elimination of multiple arcs. Let us denote the digraph obtained this way: $G_1 = (\mathcal{C}, A_1)$.

The second step of the preprocess is to construct the transitive closure of $G_1$, denoted: $G_1^\triangleright = (\mathcal{C}, A_1^\triangleright)$. In our case, $G_1^\triangleright$ consists of 814 nodes and 17,782 arcs! Nevertheless, it can be handled without problems by both Pajek and Gephi.
DEBUGGING THE DIGRAPH (LOOKING FOR LOOPS AND CYCLES): VERIFICATION OF CURRICULUM DEVELOPMENTS

Once the digraph $G_1^c$ is obtained, it is necessary to debug it. The objective is to detect possible inconsistencies in the contents sequencing. These inconsistencies arise when contents that are prerequisites of themselves are found, that is, when there are loops in the digraph $G_1^c$ associated to the prerequisite relation.

Thanks to the computational approach to the digraph $G_1^c$ developed (using Pajek and Gephi), six loops have been detected. Figure 3 shows a graphic representation of the digraph. There are so many arcs that it is almost opaque. Nevertheless, the six loops can be seen in the upper left and right corners. They are more clearly seen in Figure 4.

![Figure 3. A graphic representation of the arcs in digraph $G_1^c$.](image)

![Figure 4. Loops in digraph $G_1^c$ (magnified).](image)

It is necessary to determine the reasons why those loops do appear in the digraph $G_1^c$. First, the origin and destination nodes of those loops are identified using Pajek. This way, 6 nodes are identified. More precisely, those nodes correspond to the contents:

- Factor de un número entero (Factor of an integer number),
- Multiplicación de números enteros (Product of integers),
- Representación gráfica de una función (Graphic representation of a function),
- Función convexa (Convex function),
A. Martínez-Zarzuelo et al.

- Función cóncava (Concave function), and
- Puntos de inflexión de una función (Inflection points of a function).

If we focus on the contents corresponding to Factor de un número entero (Factor of an integer number) and Multiplicación de números enteros (Product of integers), and the subgraph of $G_1^c$ containing just these two nodes, the subgraph has 4 arcs, two of which form a directed cycle and the two others are loops (Figure 5).

**Figure 5.** Two loops in digraph $G_1^c$ (factors of an integer number & product of integers).

Remark: In this and the following figures, the contents appear in Spanish, as they are software output captures.

If we check the digraph $G_1$, these two arcs forming a cycle already appear there, so that is the reason for finding these loops in $G_1^c$.

But digraph $G_1$ should be antisymmetric, so no loops should arise. In this case the software has detected an erratum in the concept selection. The concept “factors of an integer number” should really be divided into “factorization of an integer number” and “factors of a product of integers”. Then “factors of a product of integers” would be a prerequisite for “product of integers” and “product of integers” would be a prerequisite for “factorization of an integer number”. This was exactly the kind of errata we expected to be able to debug.

Following the same procedure, the subgraph of $G_1^c$ containing the contents Representación gráfica de una función (Graphic representation of a function), Función convexa (Convex function), Función cóncava (Concave function) and Puntos de inflexión de una función (Inflection points of a function), is represented in Figure 6.

**Figure 6.** The other loops in digraph $G_1^c$. 
In Figure 6 it is clear that there are two arcs between each pair of nodes, so the subgraph is not antisymmetric. Again, these cycles already appear in digraph $G_1$. The corresponding subgraph of $G_1$ can be found in Figure 7.

Two directed cycles are detected in digraph $G_1$, one involving: Representación gráfica de una función, Función convexa y Puntos de inflexión de una función (Graphic representation of a function, Convex function, Inflection points of a function); and another one involving: Representación gráfica de una función, Función cóncava y Puntos de inflexión de una función (Graphic representation of a function, Concave function, Inflection points of a function).

This case is really interesting. It reveals a finding of the software regarding concepts: what comes first? The concept “graphic representation of a function” or the concept “concavity and convexity”? Probably the concept “graphic representation of a function” needs to be split into two different concepts:

- “intuitive ideas regarding graphic representation of a function and plotting (in the sense of computers, by representing many points)”
- “formal graphic representation of a function (based on the determination of local extrema and inflexion points)”.

Therefore the system has discovered a detail that we, that considered that mastered the subject, had not taken into account!

The situation is similar to rule based expert systems verification (Laita & Ledesma, 1997; Laita & Roanes-Lozano, 1999): an error in the knowledge of the expert cannot be detected at the verification step, but logical contradictions can be detected.
CONCLUSIONS

Techniques borrowed from graph theory and the use of networks analysis software have allowed us to propose a methodology for debugging curriculum developments. This methodology has been developed and validated in the field of mathematics. More precisely, it has been applied to the mathematical knowledge of the education stage “E.S.O.” (compulsory Secondary Education) of the Spanish educational system.

The peculiar internal logic structure of this subject has allowed determining an epistemological criterion of knowledge structuring based on meaningful learning. A binary relation has been formally defined on this basis. Defining this binary relation has allowed us establishing a concise sequencing of educative contents. Moreover, starting from graphic representations techniques of knowledge such as conceptual maps and pathfinder associative networks, the graph structure has turned out to be ideal for modeling contents and the binary relation between them.

Using software has made possible to deal with a big amount of information. The software used, specialized in network analysis, has been key for analyzing the digraph created. It has been necessary to apply some preprocesses to make it practical (such as eliminating multiple arcs or computing the transitive closure).

Finally, thanks to the use of graph theory techniques and this specialized software, loops have been found in the graph and have been analyzed. These loops, interpreted as the existence of contents that are prerequisite of themselves, lead to interesting findings.

Two loops derive from an erratum: certain contents should be split into two different ones.

The other loops lead to an unexpected and far more interesting finding: the convenience of refining a standard concept (“graphic representation of a function”) and distinguishing a concept similar to “manual plotting of a function” from “formal graphic representation of a function”.

This kind of findings and reflections are contributions to the processes of sequencing contents and, therefore, to the processes of planning teaching and instructional design.

NOTES

1 http://vlado.fmf.uni-lj.si/pub/networks/pajek/
2 http://gephi.github.io/

ACKNOWLEDGEMENTS

This work was partially supported by the research projects TIN2015-66471-P and EDU2009-14773-C02-01 (Gobierno de España) and S2013/ICE-2845 (CASI-CAM) (Comunidad de Madrid).

REFERENCES


Real decreto por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria (Real Decreto 1007/1991, de 14 de junio). *Boletín Oficial del Estado, nº 152, 1991, 26 de junio*.

Real decreto por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria (Real Decreto 1631/2006, de 29 de diciembre). *Boletín Oficial del Estado, nº 5, 2007, 5 de enero*.

Real decreto por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato (Real Decreto 1105/2014, de 26 de diciembre). *Boletín Oficial del Estado, nº 3, 2015, 3 de enero*.


A. Martínez-Zarzuelo et al.


★★★★★