Concept Maps: Making Learning Meaningful
Proc. of Fourth Int. Conference on Concept Mapping
J.Sánchez, A.J.Cañas, J.D.Novak, Eds.
Viña del Mar, Chile, 2010

A TAXONOMIC SCHEME FOR PROPOSITIONAL ANALYSIS

Jerson Geraldo Romano Jr & Paulo Rogério Miranda Correia, Universidade de São Paulo, Brasil

Email: prmc@usp.br

Abstract. Propositions are the concept maps' building blocks. Despite prepared to answer the same focal question, different authors can produce highly unique concept maps (Cmaps). We propose a taxonomic scheme for classifying propositions into 6 categories to explore in depth the latent information we can obtain comparing a set of Cmaps. We analyzed 55 Cmaps and the results were obtained by using descriptive statistics and hierarchical cluster analysis (multivariate exploratory approach). We characterized 6 clusters containing Cmaps with distinguishable features, which helped to assess the students' understanding about the relationships among science, technology, and society (STS). Cmaps rich in dynamic causative propositions indicated a sophisticated understanding about the STS relationship.

1 Introduction

Concept mapping is a well-established technique that allows knowledge to be graphically represented (Novak, 2010). It has been widely used for teaching and corporate purposes, with a broad range of goals: assessing prior knowledge, eliciting, archiving and sharing expert knowledge, and fostering collaboration are just a few of the practical and educational uses for this technique (Fourie & Westhuizen, 2008; Kyrö & Niskanen, 2008; Novak, 2010; Torres & Marriott, 2009).

Concept maps (Cmaps) can be defined as a set of concepts into a propositional network. Propositions are called semantic units, and they are critical to make clear the mapper's conceptual understanding. The literature has presented some interesting works focused on propositional analysis (Safayeni, Derbentseva, & Cañas, 2005; Miller & Cañas, 2008; Derbentseva, Safayeni, & Cañas, 2007). The combination of the available information has allowed us to improve our capacity to explain the propositions that are present in Cmaps. Static, dynamic causative, and dynamic non-causative are the main categories used to describe propositions (Safayeni, Derbentseva, & Cañas, 2005; Miller & Cañas, 2008; Derbentseva, Safayeni, & Cañas, 2007). Our experience with propositional analysis suggests the inclusion of a parameter to consider how dynamic propositions show the interdependence between the concepts: implicitly, partially explicit, and totally explicit. A taxonomic scheme for classifying propositions is presented in Figure 1.

The aim of this work was to evaluate the use of the proposed taxonomy to classify the propositions (n=825) of Cmaps (n=55) about the relationships among science, technology and society (STS). The STS approach is necessary to change science teaching throughout the formal education, and it should be implemented by using innovative instructional strategies (Bybee & Fuchs, 2006; Correia et al, 2010; DeHaan, 2005; Donnelly, 2004; Holbrook & Rannikmae, 2007; Infante-Malachias & Correia, 2007; Santos, 2007).

2 Research procedures

2.1 Setting and data collection

The Cmaps (n=55) considered in this work were obtained during the ACH 0011 Natural Sciences course, which is offered for all first-year students at Escola de Artes, Ciências e Humanidades (School of Arts, Science and Humanities at São Paulo University). The main goal of this course is to provide a comprehensive view of the impact caused by scientific and technological development in our society (Correia et al, 2010). Sixty students from four different undergraduate courses were grouped in one classroom, for two-hour weekly classes over a period of fifteen weeks.

Considering the introductory scope of Natural Sciences course and the diverse audience, the challenge to satisfy students' expectations required innovative methodological strategies.

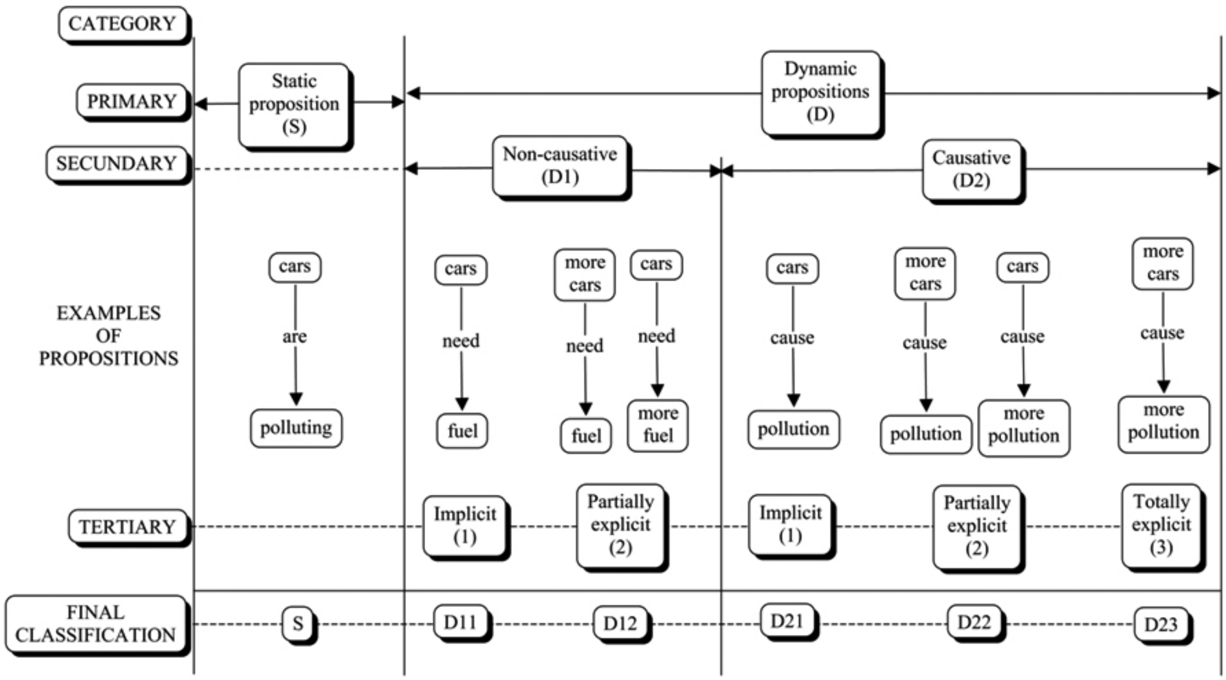

Figure 1. Taxonomic scheme for classifying propositions. Primary categories: static (S) and dynamic (D); secondary categories: non-causative (D1) and causative (D2); tertiary categories: implicit (1), partially explicit (2), and totally explicit (3).

Students were introduced to concept mapping and trained to elaborate both manuscript and digital Cmaps using the Cmaptools software at the beginning of the ACH 0011 course (Correia, Infante-Malachias, & Godoy, 2008). The aim was to ensure a reasonable training session for students in order to avoid the inappropriate use of concept mapping, as highlighted in the literature (Cañas & Novak, 2006; Correia, Infante-Malachias, & Godoy, 2008).

Bioethics was the main issue discussed at the end of the ACH 0011 course (classes 11-15). Table 1 summarizes part of the course schedule, showing when discussions about bioethics took place and the assigned activities using Cmaps. The conceptual framework explored in these classes involved evolution and the origin of life (class 11), DNA and molecular biology (class 12), medical implications of molecular biology (class 13), and the Brazilian biosecurity law (class 14). The final course test (class 15) involved the individual construction of a half-structured concept map (HSCmap) with 9 concepts to address the following focal question: "How does bioethics regulate the relationship between science and society?". The concepts "more controversy" and "more technology" were imposed to the students, and they must be present in their HSCmaps (Figure 2). Both the how-type focal question and the quantified concepts are strategies to promote dynamic, rather than static propositions (Derbentseva, Safayeni, & Cañas, 2007).

Class #	Pedagogical activities	Cmap activities
11	Preparatory reading assigned *"Explaining the very improbable" (Dawkins, 1996)*	Individual Cmaps about the preparation reading (before the class) Collaborative Cmaps (after the classroom discussions)
12	Preparatory reading assigned *"Science, Genetics and Ethics: Memo for Tony Blair" (Dawkins, 2004)*	Individual Cmaps about the preparation reading (before the class) Collaborative Cmaps (after the classroom discussions)
13	Film session: episode 1 of *"DNA, the promise and the price" (Discovery Channel)*	-
14	Brazilian Biosecurity Law	Collaborative Cmaps about the discussion during the classes 13-14
15	ACH 0011 course final test	Half-sctructured concept map (HSCmap) to answer "How does bioethics regulate the relationship between science and society?" Obligatory concepts: more controversy and more technology

Table 1. Pedagogical organization of the activities about bioethics during ACH 0011, Natural Sciences course.

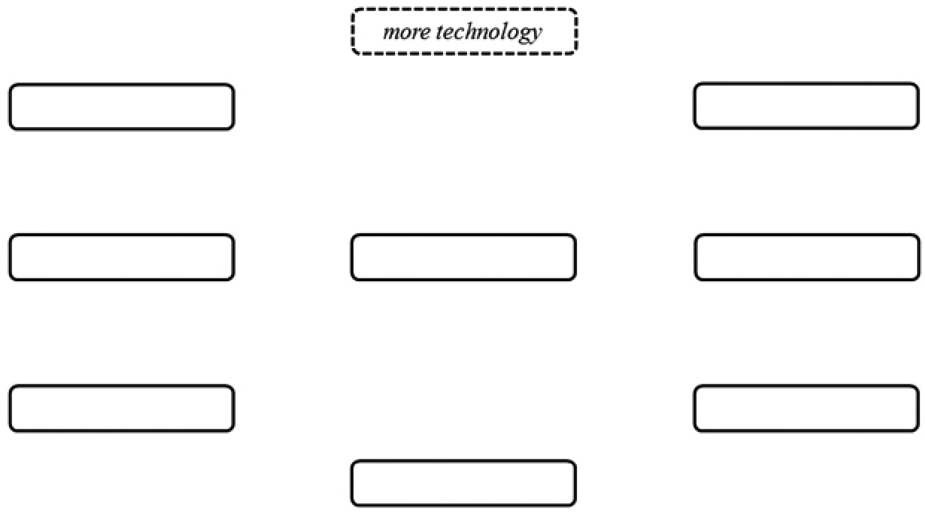

Figure 2. HSCmap with 9 concepts used in the final test of the ACH 0011 course. The concepts “more controversy” and “more technology” were obligatory, and the latter was positioned as the starting concept of the HSCmaps.

2.2 Propositional analysis of HSCmaps

Several works dealing with rubrics for evaluating the semantic features of propositions were recently proposed (Derbentseva, Safayeni, & Cañas, 2007; Miller & Cañas, 2008; Safayeni, Derbentseva, & Cañas, 2005). Inspired by those ideas, we devised the taxonomic scheme presented in Figure 1. The analytical method to classify each proposition into one of the proposed categories (S, D11, D12, D21, D22, and D23) is described in Table 2. A 4-question sequence was enough to make an appropriate choice.

Focus on	Category	Questions to classify the propositions	Response	Action	Result
LINKING PHRASE	Primary	#1. Does the linking phrase have a verb?	Yes	See question #2	-
			No	Static proposition	S
		#2. Is the verb an auxiliary one *(e.g. be, do , have)*?	Yes	Static proposition	S
			No	Dynamic proposition See question #3	D
CONCEPTS	Secondary	#3. Do both initial and final concepts indicate cause and consequence, respectively?	Yes	Dynamic, causative proposition See question #4	D2
			No	Dynamic, non-causative proposition See question #4	D1
	Tertiary	#4. Are the concepts quantified *(e.g. more cars, many cars)*?	Both concepts	Totally explicit (3)	D23
			Only one concept	Partially explicit (2)	D12, D22
			None	Implicit (1)	D11, D21

Table 2. Procedure for classifying propositions into the taxonomic scheme categories.

2.3 Data analysis

2.3.1 Descriptive statistics

Box-plot graphs were used to show the statistical parameters (average, median, lower quartile, upper quartile, sample minimum, and sample maximum) obtained to describe each variable considered in our analysis (S, D11, D12, D21, D22, and D23). These graphs allow checking the central tendency, the symmetry, and the dispersion of a data set at a glance (Cohen & Lea, 2004).

2.3.2 Cluster analysis: multivariate statistics

A data matrix X(55,6) containing all data from the propositional analysis was used to carry out hierarchical cluster analysis (HCA). HCA is an exploratory statistical method to classify objects (concept maps) into groups according to their similarities (Kaufman & Rousseeuw, 2005). All variables (n=6) are taken into account and this multivariate approach fits well to unveil latent information. Statistica 8 (StatSoft, Tulsa, OK, USA) was the software chosen to perform this analysis. The graphical output of HCA (dendrogram) will be used to discuss the results in the next section.

3 Results and discussion

3.1 Propositional analysis at a glance: frequency distribution to the devised categories

The 55 HSCmaps presented 825 propositions that were classified as indicated in Table 3. More than 80% of all propositions were classified as dynamic, and causative propositions (43.1 %) were slightly more frequent than non-causative (39.7 %). These results suggest the strategies adopted to promote dynamic rather than static propositions were effective. The propositions using the obligatory concepts were removed to evaluate the impact of the how-type focal question. The subset A was obtained after removing all propositions containing “more technology”, and the subset B was obtained after removing all propositions containing both “more technology” and “more controversy”. The comparison of subsets A and B with the full set of propositions showed a decrease in the proportion of D12, D22 and D23, while the amount of static propositions, D11 and D21 increased. The 3:1 ratio for dynamic (74.5%): static (25.5%) propositions observed in subset B was primarily due to the use of a how-type focal question (Table 3).

Proposition sets	Static propositions	Dynamic propositions				
		D11	D12	D21	D22	D23
Full (n=825)	142 (17.2%)	261 (31.6%)	67 (8.1%)	97 (11.8%)	191 (23.2%)	67 (8.1%)
A (n=680)	140 (20.6%)	260 (38.4%)	34 (5.0%)	97 (14.3%)	126 (18.5%)	22 (3.2%)
B (n=548)	140 (25.5%)	260 (47.4%)	13 (2.4%)	97 (17.7%)	33 (6.0%)	5 (0.9%)

Table 3. Frequency distribution of the analyzed propositions into the proposed categories.

3.2 Box-plots for the devised categories

Figure 3 presents the box-plots obtained for each category proposed in this work. The average (Av) and the median (Me) of static propositions (S) were 3±2 and 2, respectively. The category D11 presented the highest values among all D-categories (Av=5±3; Me=4), while the categories D12 (Av=1±1; Me=1), D21 (Av=2±2; Me=1), and D23 (Av=1±1; Me=1) presented lower values than S. The category D22 (Av=3±2; Me=3) presented similar values in comparison with S.

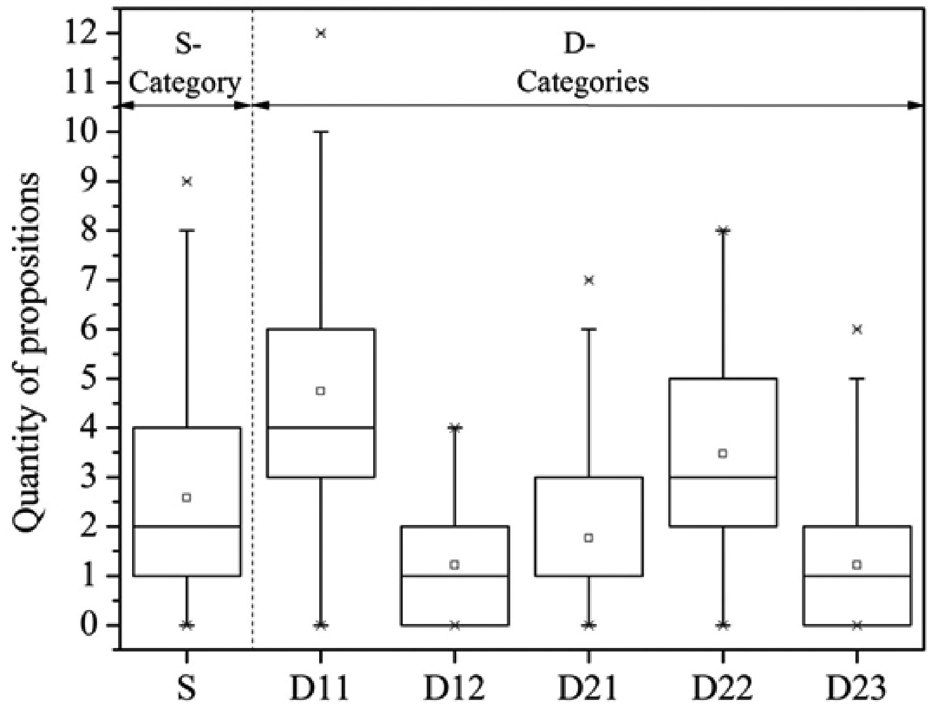

Figure 3. Box-plot obtained for each category of the propositional analysis under investigation. S-Category refers to static propositions; D-Categories refer to dynamic propositions.

3.3 Finding patterns through cluster analysis

Hierarchical cluster analysis (HCA) was used as a multivariate exploratory analysis (Kaufman & Rousseeuw, 2005). The entire data set was organized into as a matrix X(55,6). Several parameters were tested in order to achieve the most significant clusters, considering our knowledge on concept mapping. Distances between objects (HSCmaps) were measured using single, complete, and Manhattan (or City-Block) methods. Distances between clusters were measured using Euclidean and Ward's methods. The dendrogram presented in Figure 4 was obtained using Manhattan and Ward's methods for measuring distances between objects and clusters, respectively. An arbitrary value corresponding to 40% of de maximum distance between 2 objects (Cmaps) was chosen to describe the clusters to be compared. The dotted line in Figure 4 helps to identify 6 clusters (I-VI).

Despite the descriptive parameters presented in Figure 3 for all Cmaps (n=55), the average number of propositions for each group (S, D11, D12, D21, D22, and D23) was calculated for each cluster. The differences between clusters can be discussed from the data presented in Table 4. The use of black dots was a choice to make easier to compare the recognizable features of each cluster. In spite of all quantitative manipulations of the data, the qualitative approach dominates when HCA is used. Therefore, the information in Table 4 is presented in a friendly graphical manner.

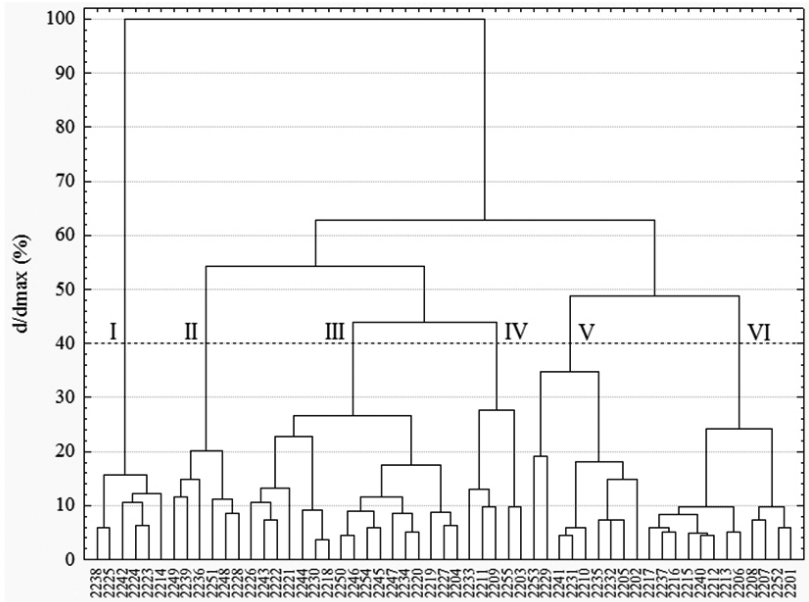

Figure 4. Dendrogram obtained by HCA considering the full data matrix X(55,6). Selected parameters for running HCA: Manhattan (or City Block) distance to measure between CMaps distance; Ward distance to measure between clusters distance. An arbitrary value of 40% of the maximum distance (dmax) was selected to compare the characteristics of the obtained clusters (I-VI).

	Cluster ID	**I**	**II**	**III**	**IV**	**V**	**VI**
	# of CMs	*6*	*6*	*17*	*5*	*9*	*12*
Categories for propositional analysis	*S*	•	•	••	••••	•	•
	D11	••••	•	••	••	•••	•••
	D12	••••	••	•••	••	•	••
	D21	••	••	••••	••	••••	•
	D22	•••	•••	•••	•••	••••	•
	D23	••	•••••	•	•••	••	•••

Table 4. Differences between clusters explained by the categories used for propositional analysis. The number of black dots indicates the relative average magnitude for each cluster. Highest values (••••), lowest values (•), intermediate values (••) or (•••).

The highest average values (••••) highlighted the differences between the clusters I (D11 and D12), II (D23), IV (S), and V (D21 and D22). The lowest average values (•) were useful to characterize the cluster VI (S, D21, and D22), and cluster III (D23). Intermediate values (••/•••) can be used to improve the description among clusters I-VI. The categories proposed in our taxonomic scheme for propositional analysis made possible to interpret the result obtained from HCA (Figure 4).

- Cluster I contains Cmaps rich in dynamic, non-causative propositions partially explicit/implicit (D11, D12).
- Cluster II contains Cmaps rich in dynamic, causative propositions totally explicit (D23).
- Clusters III and V contain Cmaps rich in dynamic, causative propositions partially explicit/implicit (D21, D22).
- Cluster IV contains Cmaps rich in static propositions (S).
- Cluster VI contains Cmaps with sorted propositions.

3.4 Illustrative concept maps

Figure 5 shows representative Cmaps to illustrate the main features presented by the clusters I, II, and IV.

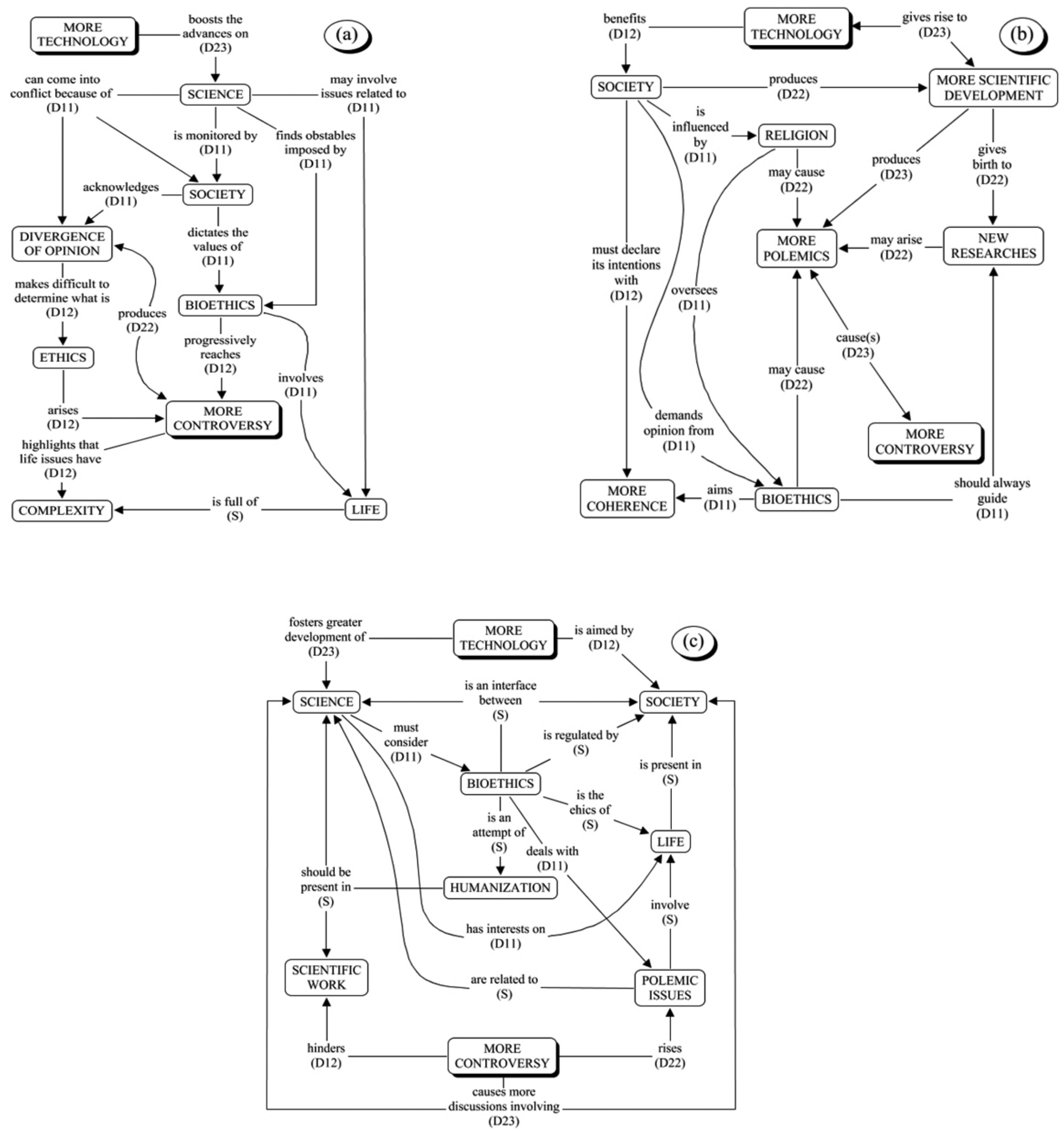

Figure 5. Selected Cmaps prepared by the students during the final course test: (a) cluster I (mainly D11 and D12 propositions), (b) cluster II (mainly D23 propositions), and (c) cluster IV (mainly S propositions).

All Cmaps in Figure 5 have a high proposition/concept ratio (Figure 5a: 1.78; Figure 5b: 1.89; Figure 5c: 2.11). The students could manage Cmaps with extremely dense networks, confirming they were well training at the beginning of the course. Most propositions in Figure 5a were classified as dynamic and non-causative, indicating the student went beyond the mere description of the concepts:

- more technology - boosts the advances on ⟶ science (D23)
- science - is monitored by ⟶ society (D11)
- science - can come into conflict because of ⟶ divergence of opinion (D11)

- society - acknowledges ⟶ divergence of opinion (D11)
- society - dictates the values of⟶ bioethics (D11)
- bioethics - progressively reaches⟶ more controversy (D12)

The cause-consequence relationship among concepts was even more obvious in Figure 5b. The author mainly used dynamic and causative propositions, and the dynamic thinking is expressed throughout the Cmap network:

- more technology - benefits - society (D12)
- more technology - gives rise to - more scientific development (D23)
- more scientific development - gives rise to - more technology (D23)
- more scientific development - gives birth to - new researches (D22)
- new researches - may arise - more polemics (D22)

Although all strategies used to promote dynamic propositions, some Cmaps presented too many static propositions. Figure 5c shows a Cmap with 10 dynamic and 9 static propositions. The concept "bioethics" has 5 static propositions, confirming the descriptive approach chosen by the author.

- bioethics - is an interface between - science (S)
- bioethics - is an interface between - society (S)
- bioethics - is regulated by - society (S)
- bioethics - is the ethics of - life (S)
- bioethics - is an attempt of -humanization (S)

4 Summary

Propositions are essential to understand the meanings from the Cmap networks. They can be thoroughly analyzed to find distinct differences [1] according to the verb used in the linking phrase (static and dynamic), [2] considering the potential cause-effect relationship between initial and final concepts (causative and non-causative), and [3] checking if the concepts are quantified (implicit, partially explicit, and totally explicit. These 3 aspects were the basis to create the proposed taxonomic scheme for classifying propositions presented in this work. Preliminary data from a set of 55 Cmaps confirmed the possibility of identifying patterns, which can be related to the authors' understanding about the subject and to their familiarity with concept mapping. Despite this work being explored Cmaps about science teaching, we believe that the taxonomic scheme for propositional analysis has a wide range of applications, including both educational and corporative uses.

5 Acknowledgements

The authors thank CAPES (3555/09-7), CNPq (553710/2006-0) and FAPESP (06/03083-0, 08/04709-6) for funding our research group. J. G. Romano Jr thanks the Secretaria Estadual da Educação for his scholarship. We also thank Profa Dra Káthia Maria Honorio for all insightful discussion about exploratory multivariate analysis and hierarchical cluster analysis. Finally, we are in debt to the students responsible for the concept maps presented in this work (Gustavo Szabo and Silva, Fig 5a; Madoka Hishi, Fig 5b; and Alison Rodrigues de Oliveira, Fig 5c).

References

Bybee, R. W., & Fuchs, B. (2006). Preparing the 21st century workforce: a new reform in science and technology education. Journal of Research in Science Teaching, 43(4), 349–352.

Cañas, A. J., & Novak, J. D. (2006). Re-examining the foundations for effective use of concept maps. In A. J. Cañas & J. D. Novak (Eds.), Proceedings of the Second International Conference on Concept Mapping (pp. 247-255). San Jose, Costa Rica: Universidad de Costa Rica.

Cohen, B. H., & Lea, R. B. (2004). Essentials of statistics for the social and behavioral sciences. Hoboken: John Wiley & Sons.

Correia P. R. M., Infante-Malachias M. E., & Godoy C. E. C. (2008). From theory to practice: the foundations for training students to make collaborative concept maps. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 146-153). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Correia P. R. M., Valle, B. X., Dazzani, M., & Infante-Malachias, M. E. (2010). The importance of scientific literacy in fostering education for sustainability: theoretical considerations and preliminary findings from a Brazilian experience. Journal of Cleaner Production, 18(7), 678–685.

Dawkins, R. (1996). The blind watchmaker: why the evidence of evolution reveals a universe without design. 2nd Ed. New York: W. W. Norton & Co. Inc.

Dawkins, R. (2004). A Devil's chaplain: reflections on hope, lies, science and love. New York: First Mariner Books.

DeHaan, R. L. (2005). The impending revolution in undergraduate science education. Journal of Science Education and Technology, 14(2), 253-269.

Derbentseva, N., Safayeni, F., & Cañas, A. J. (2007). Concept maps: experiments on dynamic thinking. Journal of Research in Science Teaching, 44(3), 448-465.

Donnelly, J. F. (2004). Humanizing science education. Science Education, 88(5), 762-784.

Fourie, L. C. H., van der Westhuizen, T. (2008). The value and use of concept maps in the alignment of strategic intent. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 399-406). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Holbrook, J., & Rannikmae, M. (2007). The nature of science education for enhancing scientific literacy. International Journal of Science Education, 29(11), 1347-1362.

Infante-Malachias, M. E., & Correia, P. R. M. (2007). Problemas complejos en el mundo postindustrial. Novedades Educativas, 203, 29–33.

Kaufman, L., & Rousseeuw, P. J. (2005). Find groups in data: an introduction to cluster analysis. Hoboken: John Wiley & Sons.

Kyrö, P., & Niskanen, V. A. (2008). A concept-map modeling approach to business planning in a computer environment. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 262-269). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Miller, N. L., & Cañas, A. J. (2008). A semantic scoring rubric for concept maps: design and reliability. In: A. J. Cañas, P. Reiska, M. Ahlberg, & J. D. Novak (Eds.), Proceedings of the Third International Conference on Concept Mapping (pp. 253-260). Tallin (Estonia) & Helsinki (Finland): OÜ Vali Press.

Novak, J. D. (2010). Learning, creating, and using knowledge: concept maps as facilitative tools in schools and corporations. 2nd Ed. New York: Routledge.

Safayeni, F., Derbentseva, N., & Cañas, A. J. (2005). A theoretical note on concepts and the need for cyclic concept maps. Journal of Research in Science Teaching, 42(7), 741-766.

Santos, W. L. P. (2007). Scientific literacy: a freirean perspective as a radical view of humanistic science education. Science Education, 93(2), 361-382.

Torres, P. L., & Marriott, R. C. V. (2009). Handbook of research on collaborative learning using concept mapping. Hershey: IGI Global.