

# Concept networks in learning and the epistemic support of their key concepts

Ismo T. Koponen and Maija Nousiainen

**Abstract** Concept maps are used in teaching and learning as representations of students' understanding of conceptual knowledge. Concept maps are basically networks of interlinked web of concepts. A long-standing problem in educational research is identifying the key concepts of importance in such networks. Here we use network analysis to examine students' representations of the relatedness of physics concepts in the form of concept maps, and suggest how key concepts and their epistemic support can be identified. The concept maps are analysed as directed and weighted networks, where nodes are concepts and links represent different types of connections between concepts. The notion of key concept is operationalised through the communicability, separately for out-going and in-coming weighted links. Using a collated concept network based on a sample of 12 original concept maps constructed by university students we show that the communicability is a simple and reliable way to identify the key concepts and examine their epistemic justification within the network.

## 1 Introduction

Concept maps are widely used as representations of knowledge in the teaching and learning sciences [13, 8, 9, 15]. Concept maps are representations of students' views of conceptual relationships, some of the connections conceptually close and some distant. Most often concepts maps are used in contexts of expressing so-called declarative knowledge, when students' conceptual knowledge can be written down

---

Ismo T. Koponen

Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland e-mail: ismo.koponen@helsinki.fi

Maija Nousiainen

Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland e-mail: maija.nousiainen@helsinki.fi

or expressed by writing in symbolical form. Consequently, concept maps are often used as tools of assessment and evaluation in learning, because they are assumed to reflect the students' conceptual understanding [17, 12, 14]. For the assessment and evaluation of concept maps there are many well-established techniques [17, 12, 14, 18]. However, these techniques (when they are quantitative) are still too limited in their scope, or (when qualitative), they are too ambiguous. The quantitative methods of concept maps analysis have traditionally approached the question of the quality of the maps by counting the hierarchical levels produced by the node-link-node connections and the number of cross-links within a given hierarchical level. These kinds of analysis methods pay attention to how concepts are interconnected locally [17, 12, 14, 18] but often fall short in not taking into account the global structure of the interlinked connections [10, 11]. Moreover, the link-counting method is biased to yield always better results for maps with greater density of links per node, in which case the global properties of concept maps remain subordinated to the density of links.

We approach the analysis of concept maps from the viewpoint of the cartography of knowledge [2, 3, 4] by using the network approach in characterizing the concept maps to identify their key concepts. The network-based approaches pay due attention to the global features of interlinked knowledge structures, thus providing promising means of augmenting traditional approaches that have been used in the analysis of students' knowledge structures and their organization [10, 11]. Here we assume that the key elements of the students declarative knowledge are those elements which have an important local and global role in the concept networks. The role of local importance is like its role in traditional quantitative analysis in that it quantifies node's connections to neighboring elements. The global role derives from global connectivity; how the element is connected through contiguous connections to other elements in the network. Network theory provides several operationalizations of such properties of which the communicability, which is based weighted counting of walks between nodes [6, 7, 5, 1]. Such communicability between the nodes is a promising approach to investigate the global connectedness within the concept network. The communicability is also very closely related to how different elements in the network provide support and justification for individual elements within it [11].

We examine here 12 students' concept maps based on the linking of elements of conceptual knowledge of electricity and magnetism, where simple rules to connect the diverse types of elements are used to construct the map [11, 16]. However, instead of analysing each map individually we analyse the collated map where all acceptable elements (121 altogether) and their connections (787 altogether) found in the 12 original maps are aggregated. We do not discuss here the content or content analysis of networks, which is presented in detail elsewhere [10, 15, 16]. Instead, we focus on assessing the local and global roles of the nodes in a network. We ask, if the role of nodes in concept network are the determined only by the local features, which justifies the traditional local link-counting methods of concept map evaluation, or whether global features are so essential that local analysis methods are bound to fail. In the latter case students' conceptual landscapes are truly com-

plex and their proper analysis and cartography requires novel approaches sensitive to global features and their complexities.

## 2 Empirical sample: Concept maps

The concept networks to be studied here are based on concept maps drawn by pre-service teachers. Our sample of 12 concept maps comes from a 7-week teacher preparation course, where topics in electricity and magnetism were discussed on the level of a first year introductory university course and also from the viewpoint of teachers' subject matter knowledge needed for purposes of designing teaching approaches for upper secondary school [10, 15]. The task was to represent how important concepts, principles and laws, and the essential experiments and models supporting their introduction, could be represented in a logical and coherent way. The students represented their views on the possible ways of introducing the new concepts in the form of a concept map and in a written report explaining the content of the map. The process of constructing a concept map consists of a few basic steps, based on a special type of selection of the concepts and links connecting the concepts. The nodes in the map can be: 1) concepts or quantities; 2) laws; 3) models or 4) experiments. The original links on the maps are simple actions or procedures, e.g., change in a quantity, setting the value of a quantity, or determining the value of a quantity [10, 15]. The concept maps thus contain the concepts students think are of importance. In all the maps there are  $N = 121$  different concepts in total, of which in average 60 appear in a given single map. The number of different links in all the maps is 787, with about 100 links in a single map. One example of a concept map constructed by a student is shown in Figure 1. The concepts, laws, principles and models that appear most often in the maps are collected and numbered in Table I. Further details of the content of the maps are not, however, important in what follows. The concept network which is analysed here is a collated network including all 121 nodes and 787 links. The collated map provides a window to the totality of the concepts and their interconnected relations which entered into the discussions during the course through the exercises and discussions connected to them. It is also suitable in demonstrating the differences between results based on traditional link counting in comparison to analysis based on the communicability.

## 3 Method of analysis

The analysis is based on a mixture of qualitative and quantitative methods. First, qualitative analysis is used in assessing the degree of epistemic justification for the knowledge expressed in the maps and written reports. Second, quantitative analysis based on the operationalisation of importance as the communicability is used to



are described correctly, so that the experiments are feasible ones and the models are rationally developed; 4) Valid justification. Justification consists of rational and logically correct argumentation so that stated claims proceed in a logically sound order. These four epistemic criteria are cumulative and thus the scoring  $S$  from 1 to 4 gives the order in which the above epistemic norms should be fulfilled and provides the simple basis to quantify the epistemic strength as  $s = S/S_{max}$ , with  $s \in [0, 1]$ . The links are simply taken as they are drawn, no evaluation of their content has been done because the linking words are not as informative as the content of nodes.

After the interpretative analysis, the strengths  $s$  of nodes carry the most important content related information of interest here for identification of the key concepts. Epistemically strong nodes have strengths  $0.75 < s < 1$ , while weak nodes have values  $0.0 < s < 0.25$ . The strength of node  $i$  is denoted by  $s_i$ , while  $w'_{ij}$  is the strength of the link from node  $i$  to node  $j$ . The information concerning epistemic support (through node strengths) is taken into account in the further analysis by means of rescaling the strengths of links. The epistemic strength of the initiating node and the link emerging from it are aggregated to form a new weight

$$w_{ij} = \beta s_i (w'_{ij})^\beta \quad (1)$$

This rescaling is motivated by the notion that directed links  $w'_{ij}$  pass supporting information from node  $i$  to node  $j$ . The prefactor  $\beta$  and power-law factor allows enhanced weighting of strong links so that for  $\beta \gg 1$  only strong links with  $w'_{ij} \approx 1$  survive while for  $\beta \ll 1$  all links are weighted equally. The re-weighted network described through the weighted adjacency matrix  $\mathbf{W}$  with elements  $[\mathbf{W}]_{rs} = w_{rs}$  is then analysed as a weighted directed network.

When the networks are formally described using the weighted adjacency matrix  $\mathbf{W}$ , it becomes possible to operationalise the notion of node importance in passing the epistemic support from one node to another in terms of the communicability of nodes which pays attention to walks between nodes. The communicability between nodes can be defined in terms of a weighted sum  $G_{pq}$  of walks between nodes  $p$  and  $q$ , where the weight is given by an inverse of a factorial of the length of the walk [6, 7, 5]

$$G_{pq} = 1 + \frac{[\mathbf{W}^1]_{pq}}{1!} + \frac{[\mathbf{W}^2]_{pq}}{2!} + \frac{[\mathbf{W}^3]_{pq}}{3!} + \dots = [e^{\mathbf{W}}]_{pq} \quad (2)$$

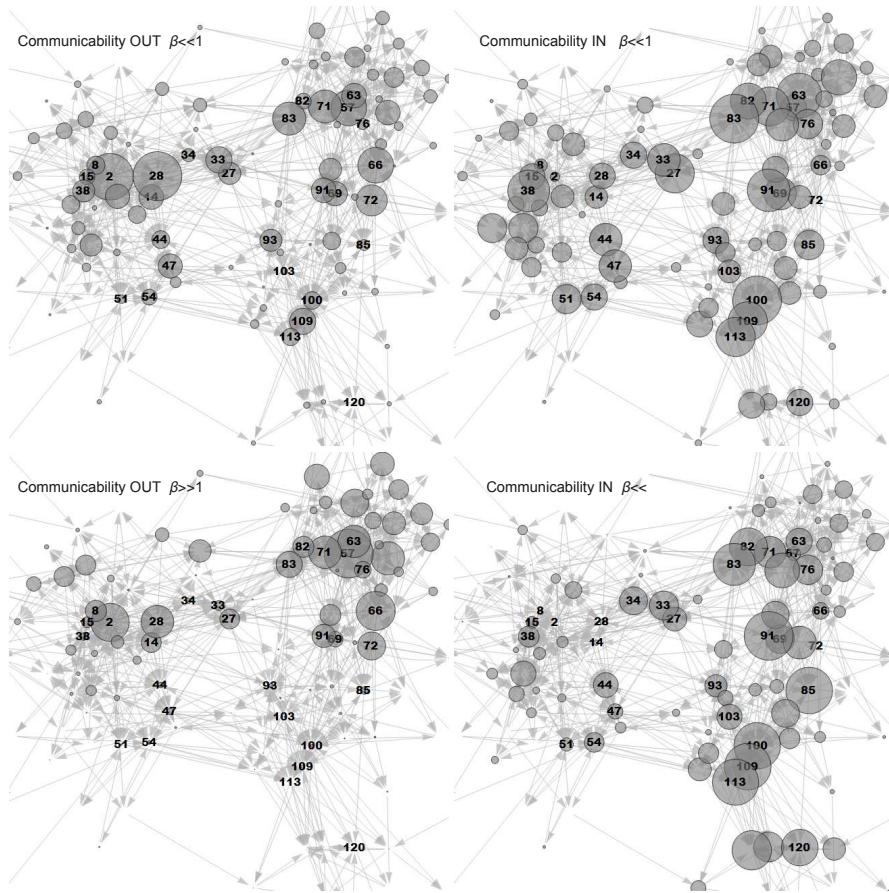
The communicability of a given node  $v$  through out-going (OUT) and in-coming links (IN) are then defined as the total out-communicability and in-communicability, respectively, as

$$G_{OUT}(v) = \sum_{p \neq v} G_{vp}, \quad G_{IN}(v) = \sum_{p \neq v} G_{pv} \quad (3)$$

The total communicability has the advantage of being a simple and robust measure closely related to the passing of information in the network. Moreover, in the limit  $\beta \ll 1$  it reduced to simple degree counting of nodes and allows direct comparison of results for directed weighted networks with unweighted link-counting based methods of concept map analysis.

## 4 Results

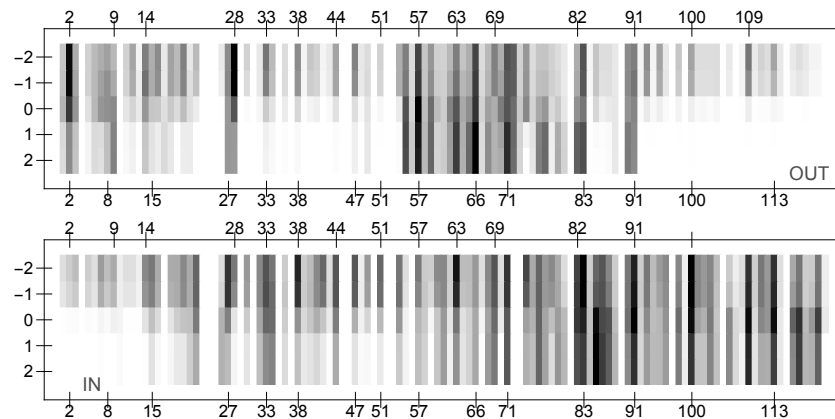
The collated concept network representing all acceptable connections from the 12 student drawn concept maps has 121 nodes and about 787 links. The concepts network with nodes scaled according to the out- and in-communicabilities  $G_{OUT}(v)$  and  $G_{IN}(v)$ , respectively, of a given node  $v$  is shown in Figure 2. The numbering of the key nodes are explained in Table I. The nodes in the network have well defined content and physical meaning [10, 15, 16]. Here, however, we are not focusing on the content of the network but rather on its structure and structural analysis that identifies the key nodes. Note that the modularity of the structure seen in Fig. 2 is a consequence of the task structures separate for tasks of electricity, magnetism and electromagnetism, seen as three different cliques in the network.



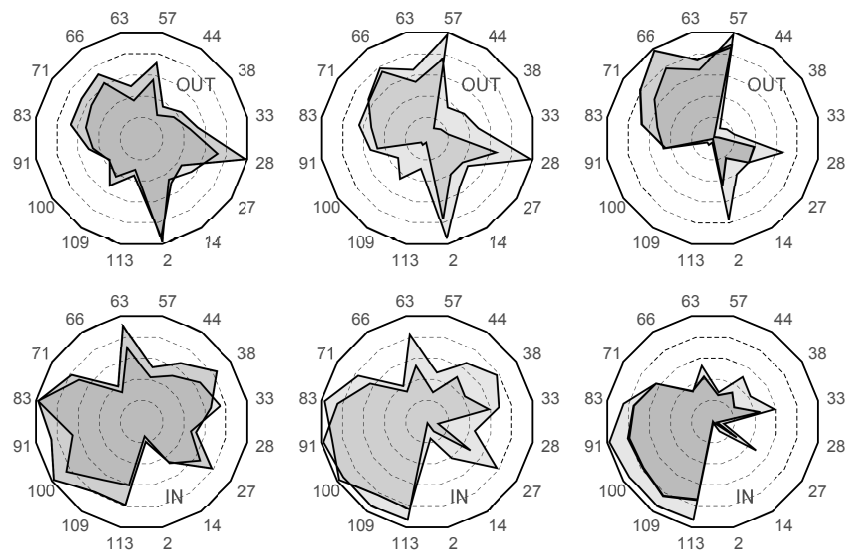
**Fig. 2** Concept maps transformed to weighted directed networks. Nodes shown in proportion to their out- and in-communicability. In the upper row node weighting is  $\beta = 10^{-2}$  and in the lower row  $\beta = 10^2$ .

Figure 2 shows that the out-communicability  $G_{OUT}(v)$  of node  $v$  through outgoing links is essentially different from its in-communicability  $G_{IN}(v)$  through incoming links. In the region  $\beta \ll 1$  where all nodes are weighted similarly the in- and out-communicabilities are directly proportional to in- and out degrees of the nodes and ranking of different nodes on the basis of the communicability is equal to the ranking obtained on the basis of in- and out-going degrees. When all the links are equally justified the network is equivalent to a network with maximally justified links, thus corresponding to experts' knowledge with its key concepts. For equally justified links, the traditional analysis methods would give rankings in complete agreement with the communicability and there is no need for more advanced analysis methods. When nodes are weighted according to their epistemic strengths, and links are rescaled accordingly as given by Eq. (1), the situation is substantially changed. Now the weighting allows us to study the epistemic strength of the conceptual knowledge as it is represented by the concept network. The lower panel of Fig. 2 shows the networks with weighting  $\beta = 10^2$  in which case only the most strongly weighted links survive and contribute to the communicability. It is now seen how only a fraction of nodes have a high communicability. In case of the out-communicability these nodes are in the cluster mediating (magnetism concepts) between the cluster of source nodes (electricity concepts) and target nodes (electromagnetism nodes), as expected. In the source cluster only a few concepts retain their role as key concepts, while in the target clusters the role of certain concepts as key concepts is enhanced. In the mediating cluster the changes are not so dramatic and many key concepts retain their roles as key concepts. Interestingly, in a different kind of analysis it is also seen that the concepts of magnetism have for students a crucial mediating role in how they organize their knowledge of electricity and magnetism [10]. The resulting set of strong nodes for  $\beta \gg 1$  represents that knowledge the student group can be assumed to justify properly and bring in their discussions in a properly justified manner; these key nodes thus provide important information of the quality of knowledge within the group and how it may enter in group discussions.

The set of key concepts and how it changes with increased weighting of epistemic strength is shown in more detail in Fig. 3 where the key concept fingerprint [11] of the network is shown. In the fingerprint the key concepts which have a high communicability are shown as black stripes, while the lesser the communicability the lighter the colour of the stripe. In the fingerprints it is shown that with increasing weighting of the epistemically strongly supported links the number of key concepts diminishes and only about 16 remain epistemically strongly supported. These sixteen nodes are: 2, 14, 27, 28, 33, 38, 44, (for electricity, names of the concepts as in Table I) and 57, 63, 66, 71, 83, 91 (for magnetism) and 100, 109 and 113 (for electromagnetism). These are the concepts students are able to justify robustly, possibly able to use, and whose significance they emphasise in the totality of the concepts of electricity. Luckily, all these concepts are also central from the point of view of content or the subject knowledge, which is not always the case for how students emphasise different concepts in their representations of knowledge [10].



**Fig. 3** Fingerprint of the key concepts in the collated concept network. The key concepts having a high communicability are shown in black. The weighting of links with  $\beta = 10^p$  changes from  $p = -2$  to  $p = 2$  as shown on the scale on the right. The communicability through outgoing links (OUT) is shown in the upper panel and through incoming links (IN) in the lower panel. Note how differently weighted networks have different sets of key nodes. The key nodes surviving the high weighting region nodes that epistemically are the most strongly supported ones.



**Fig. 4** The radar-map representation of 16 the most important key concepts in the collated concept network. In the left column, the weighting of links is  $\beta = 10^{-2}$  and compared with the degree of nodes (the largest outmost polygons). In the middle,  $\beta = 10^{-1}$  (the outmost polygon) and  $10^0$  are compared, and in the right,  $\beta = 10^1$  (the outmost polygon) and  $10^2$  are compared. The out- and in-communicability through outgoing is shown in the upper row and through incoming links in the lower row.

Another way to represent the 16 key concepts and how their importance changes with weighting is to use a so-called radar-map (or spider-web map), where each



concept is shown as a corner of a polygon. Such a radar map is shown in Fig. 4. In the left panel is shown the key concepts for  $\beta \gg 1$  in which case all the links are treated as being equally weighted and the relative strength of the communicability is similar to the case when all links would receive weight 1. In this case, as Fig. 4 shows, the (relative) in- and out-communicability are equal to (relative) in- and out-degrees, in which case also the importance rankings are similar. This case, in fact, also corresponds to the case of expert's knowledge when all the links are maximally well epistemically justified. When the weighting increase, the radar maps clearly show how certain concepts lose their role as key concepts and only a smaller subset survives. The last column shows the key concepts for  $\beta = 10^2$  in which case the situation remains the same with increasing  $\beta$ . In the case of weighted links, i.e. when epistemic justification is of importance, methods based on link counting are no more appropriate and would lead to unwarranted conclusions about which nodes are important. The method suitable for analysis of directed weighted networks can be based on weighted path counting and the importance of a node can then be quite naturally operationalized as the communicability of the node.

## 5 Discussion and conclusions

We have here analysed a concept network which agglomerates students conceptions of the key elements in electricity and magnetism. The network can be taken as representation of students shared conceptual landscape of these topics and including the key elements and their relationships on which the students focused their attention. The purpose of the study is to show how the communicability of nodes can be used in analysis of such a network and in revealing how different elements of the network are supported through their weighted connection to other nodes in the network. The methods to analyse such networks introduced here augment the traditional methods, which focus either on local properties through counting links [17, 12, 14]. It is shown how such methods can be augmented while at the same time retaining the direct connection to them by developing methods informed by complex network based methods of analysis. Such methods provide a selection of tools and approaches where desired properties of interest can be operationalized and made measurable in the networks [11].

The concept network analysed here represents the students understanding of physics concepts and their interrelationships and consists of 121 nodes and 787 links. When epistemic weighting of the links is taken into account the network cannot be analysed locally through its nodes' degree distribution only. Instead, global connectivity is fundamentally a governing property. Many nodes, irrespective of their degree play a significant role in connecting different regions of the network and thus hold important transmitting roles in the network. Some other nodes, however, are more significant in this role than expected on basis of their degree, while some other nodes are substantially less significant as expected on basis of their degree. In particular, the difference between out-going and in-coming links has an important

role in how epistemic support is transmitted from one node to another. The nodes that have a key role in transmitting the information can be found by focusing attention on those nodes that have a high communicability either through out-going or in-coming links. These nodes are the key concepts in the network. Such nodes include some certainly central concept like electric charge, electric field and magnetic field and mechanical work and the superposition of fields. The traditional analysis methods, however, would have missed many of these concepts and paid attention also on concepts marginally central from the point of view of content.

## References

1. Benzi, M., Klymko, C.: Total communicability as centrality measure. *Journal of Complex Networks* (1), 124–149 (2013)
2. Börner, K.: *Atlas of knowledge: Anyone can map*. MIT Press, Cambridge, MA (2015)
3. Börner, K., Scharnhorst, A.: Visual conceptualizations and models of science. *Journal of Informetrics* (3), 161–172 (2009)
4. Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., Pellegrino, D.: Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics* (3), 191–209 (2009)
5. Estrada, E.: *The structure of complex networks*. Oxford University Press, Oxford (2012)
6. Estrada, E., Hatano, N.: Communicability in complex networks. *Physical Review E* (77), 036,111 (2008)
7. Estrada, E., Hatano, N., Benzi, M.: The physics of communicability in complex networks. *Physics Reports* (514), 89–119 (2012)
8. Ingec, S.: Analysing concept maps as an assessment tool in teaching physics and comparison with the achievement tests. *International Journal of Science Education* (31), 1897–1915 (2009)
9. Kinchin, I., Hay, D., Adams, A.: How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research* (42), 43–57 (2000)
10. Koponen, I.T., Nousiainen, M.: Pre-service physics teachers understanding of the relational structure of physics concepts: Organising subject contents for purposes of teaching. *International Journal of Science and Mathematics Education* (11), 325–357 (2013)
11. Koponen, I.T., Nousiainen, M.: Concept networks in learning: Finding key concepts in learners' representations of the interlinked structure of scientific knowledge. *Journal of Complex Networks* (2), 187–202 (2014)
12. McClure, J.R., Sonak, B., Suen, H.K.: Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching* (36), 475–492 (1999)
13. Nesbit, J.C., Adesope, O.O.: Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research* (76), 413–448 (2006)
14. Nicoll, G., Francisco, J., Nakhleh, M.: A three-tier system for assessing concept map links: A methodological study. *International Journal in Science Education* (23), 863–875 (2001)
15. Nousiainen, M.: Coherence of pre-service physics teachers views of the relatedness of physics concepts. *Science & Education* (22), 505–525 (2013)
16. Nousiainen, M., Koponen, I.T.: Pre-service physics teachers content knowledge of electric and magnetic field concepts: Conceptual facets and their balance. *European Journal of Science and Mathematics Education* (5), 74–90 (2017)
17. Ruiz-Primo, M.A., Shavelson, R.: Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching* (33), 569–600 (1996)
18. vanZelev, E., Lenaerts, J., Wieme, W.: Improving the usefulness of concept maps as a research tool for science education. *International Journal of Science Education* (26), 1043–1064 (2004)