Concept Landscapes: Aggregating Concept Maps for Analysis

Andreas Mühling
Kiel University
andreas.muehling@informatik.uni-kiel.de

This article presents *concept landscapes* - a novel way of investigating the state and development of knowledge structures in groups of persons using concept maps. Instead of focusing on the assessment and evaluation of single maps, the data of many persons is aggregated, and data mining approaches are used in analysis. New insights into the "shared" knowledge of groups of learners are possible in this way. Electronic collection of concept maps makes it feasible to aggregate the data of a large group of persons, which in turn favors a data mining approach to analysis.

The educational theories underlying the approach, the definition of concept landscapes, and accompanying analysis methods are presented. Cluster analysis and Pathfinder networks are used on the aggregated data, allowing new insights into the structural configuration of learners' knowledge. Two real-world research projects serve as case studies for experimental results. The data structures and analysis methods necessary for working with concept landscapes have been implemented in the freely available GNU R package *CoMaTo*.

1. Introduction

This article is concerned with analyzing the shared knowledge structures of groups of persons with regard to their common elements as well as their inherent subjective differences. Investigating knowledge remains relevant in typical teaching scenarios: To become competent in an area, a person must - in general - acquire a certain set of skills as well as a certain body of knowledge. "Thus, one goal of instruction should be to help students acquire expert-like knowledge structures in their domain of study" (Trumpower and Goldsmith, 2004, p. 427). In some areas (e.g. sports or arts) experts are typically characterized by their skills, first and foremost. In other areas (including STEM), however, an expert must necessarily have acquired a rich and highly connected body of conceptual and factual knowledge (cf. Kinnebrew et al.; Trumpower et al., 2014; 2010). Even basic programming skills, for example, require factual knowledge about syntax elements and conceptual knowledge about program flow, among others.

Modern constructivist teaching implies that such a body of structured knowledge cannot passively be transported from teachers to learners. Instead, learning is a subjective, personal process, and teaching is mostly a fostering of learning. As Cañas and Novak (2006, p.495) put it, "[l]earning is highly idiosyncratic and progresses over time." Nevertheless, the knowledge organization of experts tends to be similar (cf. Trumpower and Goldsmith, 2004), and when teaching in larger groups, the effects of teaching will most probably tend to affect many learners in similar ways. Therefore, investigating how knowledge develops for a whole group of learners is an insightful research endeavor as well: "An excellent way of understanding the mental world of an individual, group or scientific community, or culture is to find out how they organize their world into concepts." (Goldstone and Kersten, 2003, p.601).

So, when investigating knowledge, either the subjectively constructed, idiosyncratic knowledge structure of a *single person* or the knowledge structures of *several persons* regarding their common elements and inherent differences may be of interest. These two facets lead to their own respective, unique insights, as a simple example of an exam taken in a school illustrates: On the one hand, one can be interested in the particular result of one student. A teacher, for example, might be interested to know that this person failed the exam, or failed multiple exams in a row. On the other hand, one can also be interested in the aggregated results for all students. The school's principal might be interested to know that almost all students failed the exam. In the first case, no insight about the exam can be gained - unless the student would usually not fail an exam. In the second case, no insight about the individual students can be gained, in general. The emphasis on single persons is concerned with *learning processes*, whereas the emphasis on many persons is concerned with *educational processes* (i.e. teaching). The same is done in other studies that are concerned with evaluating educational processes like the international large-scale study PISA (OECD, 2012), in which measurements of many individuals are combined for analysis with the goal of deriving more information about the educational process than about any single individual taking part in the experiment.

Trumpower et al. (2010, p. 6) describe the process of analyzing knowledge as consisting of three phases: "1) knowledge elicitation, 2) knowledge representation, and 3) knowledge evaluation." The knowledge of a person is not directly observable. It is indirectly observable however by using some form of externalization. This form can be rather direct, like conducting an interview or indirect, by having the person take a multiple choice test, for example. Also, when analyzing knowledge, it must not only be externalized but also represented in a suitable form for this analysis. There will usually be very many ways to externalize and represent the knowledge; any form will have its advantages and drawbacks.

Concept maps have been established as a method of investigating the structural knowledge of a person. Based on this, we here present *concept landscapes* as a general notion for aggregating the data of multiple concept maps with the goal of analyzing this combination instead of the single maps. This allows using concept mapping for the investigation of educational processes. The contribution of this article is twofold: First, we present a framework and formal definition for the aggregation of concept maps that can be used in a broad variety of experimental setups. Second, we present methods of analysis that can readily be applied (or modified to be applied) to concept landscapes. The case studies present the results of applying these methods to real world data and show that interesting insights into educational processes can be gained in this way.

While the externalization of knowledge and its analysis don't require computer support per se, the use of data mining techniques and computer-supported analysis allows the processing of larger bodies of data. Since such large amounts of data cannot be processed manually in a reasonable timeframe, computer support is more than a way of effectively saving time: It allows new insights into the (often statistical) properties of the data. Using larger amounts of data can result in new findings that cannot be found in small samples due to noise or statistical insignificance. The proposed methods can all readily be implemented in software. For the most part, using a computer (e.g. for drawing concept maps) is a natural extension of the task and allows greatly enlarging the amount of data that can be handled. In this way, it is practical to collect and immediately analyze data from many students worldwide without additional personal labor costs. The analysis methods presented here have - among others - been implemented as part of the freely available GNU R (R Core Team, 2013) package *CoMaTo*.

The rest of the article is organized as follows: The next section presents a brief overview of the foundations of psychology and educational research on which our approach is based. Then, we present the notion of concept landscapes as well as the relevant related work. Based on this, we present how concept landscapes can be analyzed, focusing on two different methods. The methods are applied to real-world data collected in three research studies. Finally, the limitations and future directions of our research are discussed.

2. Theoretical Background

2.1. Knowledge and Learning

The kind of research we are undertaking is fundamentally tied to the organization of knowledge, its creation (i.e. learning), and the elicitation of existing knowledge. The knowledge we are interested in - subject-matter knowledge about some domain - resides in long-term memory. More specifically, in *declarative memory* that Squire (1987, p. 152) describes rather succinctly: "Declarative memory is memory that is directly accessible to conscious recollection. It can be declared. It deals with the facts and data that are acquired through learning." There are other classifications of the components of human memory as well. For example, Trumpower and Goldsmith (2004) distinguish between *propositional* and *configural knowledge*, the former referring to facts and the latter meaning the interconnected structure of knowledge. The authors de Jong and Ferguson-Hessler (1996, p. 107) use the term *conceptual knowledge*, defined as "static knowledge about facts, concepts and principles that apply within a certain domain."

In general, the organization of concepts in memory is assumed to be pivotal for the quality of a person's knowledge: "[K]nowledge requires not only acquiring facts, procedures, and concepts, but also having an understanding of the interrelationships among those facts, procedures, and concepts" (Trumpower et al., 2010). Without a structured connection to others, a concept will not be kept in long-term memory (cf. Sousa, 2009). "In general, structured knowledge enables inference capabilities, assists in the elaboration of new information, and enhances retrieval. It provides potential links between stored knowledge and incoming information, which facilitate learning and problem-solving" (Glaser and Bassok, 1989, p. 648). Trumpower et al. (2010) also state that "experts possess more knowledge and, perhaps more importantly, better organize knowledge than novices". Ruiz-Primo and Shavelson (1996, p. 570) point out:

> "Most cognitive theories share the assumption that concept interrelatedness is an essential property of knowledge. [...] As expertise in a domain is attained through learning, training and/or experience, the elements of knowledge become increasingly interconnected. [...] Assuming that knowledge within a content domain is organized around central concepts, to be knowledgeable in the domain thus includes having a highly integrated structure among these concepts."

Equally important for knowledge organization is how knowledge is created in the mind in the first place. The theory of Constructivism holds that "[k]nowledge [...] cannot be imposed or transferred intact from the mind of one knower to the mind of another" (Karagiorgi and Symeou, 2005). Fundamentally, Constructivism is built upon two principles (von Glasersfeld, 1989):

1. "knowledge is not passively received but actively built up by the cognizing subject" and
2. "the function of cognition is adaptive and serves the organization of the experiential world, not the discovery of ontological reality."

Research has offered enough insight into the neurological mechanisms of learning, that the constructivist theory can justifiably be seen as an explanatory model of human learning and knowledge construction. Sabitzer (2011) provides insights into the neurological aspects of teaching computer science, stating that learning is the process of subjectively constructing knowledge. Wittrock (1992, p. 536) states from the perspective of cognitive research: "[L]earning is not the internalization of information given to us whole by experience or analyzed by a teacher. Instead, learning consists of the active generation of meaning, not the passive recording of information.".

When trying to probe into a person's knowledge - especially given the implications of Constructivism - it is paramount to acknowledge that the object of interest - the mental model - is by itself unobservable (cf. Cooke, 1994). Instead, only a model of this model can be observed (cf. Shaw and Woodward, 1990). Each method of elicitation and each method of representation is making assumptions about the knowledge itself (cf. Shaw and Woodward, 1990). The process of externalization is also subject to uncertainties. The influences that incur these uncertainties are numerous, like problems in communication or the amount to which the knowledge that is to be externalized is compiled and more (cf. Cooke, 1994); also personal variables, like the degree of introversion have been found to influence the process (cf. Hoffman et al., 1995). While all methods inherently suffer from these uncertainties, the specific influences and their extent are dependent on the method and context. The same goes for the requirements on the elicitor (e.g. to be a subject-matter expert) or the analysis methods that are suitable afterward (cf. Cooke, 1994).

2.2. Concept Maps

Historically, *concept maps* were invented in the 1970s as a tool to help with structuring and visualizing the responses of children in clinical interviews (cf. Novak and Cañas, 2010). Later on, the use of concept maps shifted from a specific technique for data analysis to a general technique for learning, teaching, and assessing structural knowledge (cf. Novak and Cañas, 2008).

Figure 1 shows an example of a concept map. It consists of labeled entities that represent a *concept*. Concepts are defined as "perceived regularities or patterns in events or objects, or records of events or objects, designated by a label" (Novak, 2010, p. 25). "When two or more concepts are related by the use of what we will call linking words, propositions are formed. These become the fundamental units of meaning stored in our cognitive structure." (Novak, 2010, p. 26). A *proposition* is composed of the two concepts and the label of the connection itself. The basic rules for creating a concept map are summarized by Hay and Kinchin (2006, p. 129) as follows:

- "The concepts that an individual deems important in illustrating their personal understanding of a topic are placed in text-boxes and arranged hierarchically on a page (so that broad and inclusive concepts are at the top and detail or illustrative example, at the bottom)."

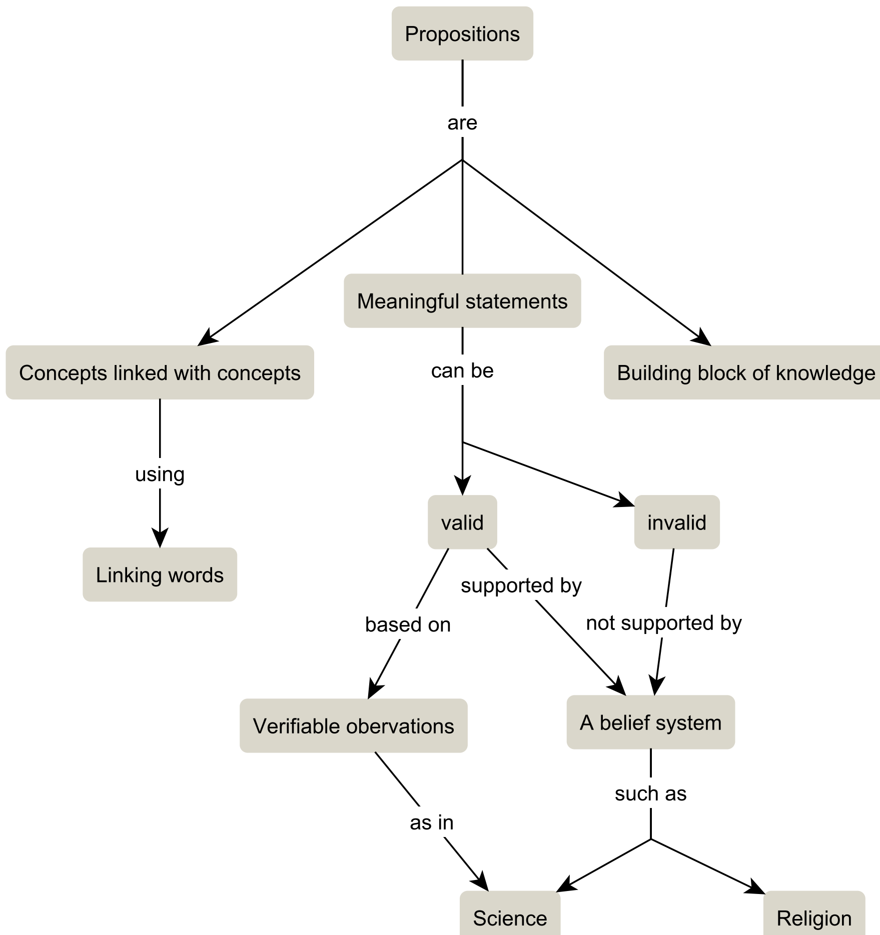


Figure 1: A concept map illustrating the concept "proposition" and its relation to knowledge. Adapted from Novak (2010)

- “Concepts are then linked with arrows that are annotated with “linking statements” to explain the nature of the link.”
- “Concepts may be listed only once, but any number of links may be made between any number of concepts at any number of conceptual links.”

The technique of concept mapping is fundamentally based on the ideas of Constructivism and meaningful learning (cf. Novak and Musonda, 1991). To this day, concept maps have been successfully used as learning and teaching aids as well as for the assessment and investigation of persons’ knowledge structures in countless scenarios, studies, and subject domains. Novak and Cañas (2010) present an in-depth review of relevant literature and many areas of application. The subject domains include human resource development (Daley et al., 2010), mathematics (Ozdemir, 2005), biology (Kinchin, 2000), training in dental medicine (Kinchin and Cabot, 2009), didactics of informatics (Gouli, 2007) computer programming (Keppens and Hay, 2008), and computer science (Sanders et al., 2008).

Most of the specific details of concept mapping as originally defined by Novak have been altered over time, allowing for a wide variety of concept mapping tasks to be found in literature - sometimes showing almost no similarity to the original (cf. Sousa, 2009). Cañas et al. (2005, p. 208) define a “well-constructed” concept map as one where “[e]ach pair of concepts, together with their joining linking phrase, can be read as an individual statement or proposition that makes sense” while “[c]oncepts and linking phrases are as short as possible, possibly single words” and “[t]he structure is hierarchical and the root node of the map is a good representative of the topic of the map.”

Following the distinctions between the different types of knowledge, concept maps can be seen as an externalization of declarative knowledge as it has been “declared” by the concept mapper (cf. Ruiz-Primo and Shavelson, 1996). Thus, a concept map, in general, is an externalization of parts of the declarative, semantic memory of a person (cf. Ruiz-Primo, 2004). Strictly speaking, concept maps are not restricted to knowledge or semantic memory, as it is also possible to externalize beliefs or personal experiences. Within these limits, however, Novak (2010) and Novak and Cañas (2010) point to the successful application of concept maps as a method for assessing or evaluating students’ knowledge structures across many different fields of study. In particular, it has been shown that assessments based on concept mapping can differentiate between the knowledge of experts and novices as well as between meaningful learning and rote learning (cf. Derbentseva et al., 2007).

The reliability and validity of concept mapping as a form of assessment have received attention from researchers over many years. Concerning the validity, the results found in the literature are generally positive: “The validity issue is relatively transparent since it is obvious that the fundamental characteristics of constructivist learning is exemplified in a well-constructed concept map” (Novak, 2010, p. 231). Rosas and Kane (2012) draw a similar conclusion. Albert and Steiner (2005) present a more detailed overview of the problem regarding validity and suggest methods of determining it. Establishing the reliability of an assessment task using concept maps is not easy, as it suffers from the fundamental problems of every method of elicitation. In general, every externalization will be influenced by many different variables which are neither completely known nor easily (or at all) measurable. The extent to which the knowledge of a person has been externalized in a concept map cannot be quantified. “[S]ources of error in a concept map test include: (a) variations in students’ concept mapping proficiency, (b) variations in the content knowledge (domain expertise) of those evaluating the concept maps, and (c) the consistency with which the concept maps are evaluated" (McClure et al., 1999, p. 477). However, at least under certain conditions, the reliability has been established by several different studies (cf. Novak, 2010).

3. Concept Landscapes

Following the motivation presented in the introduction, our goal is to aggregate concept map data *systematically* to gain insights into the state and development of knowledge of a *group of persons*. Specifically, we are trying to provide a framework of definitions that makes working with aggregated concept maps easy. We have chosen concept maps as our input data because of their long history in assessments of knowledge. The method is especially useful for digitally created concept maps, which can directly be used as input data for the data mining processes. Kwon and Cifuentes (2009) investigated the difference between digitally drawn concept maps and pen and paper based ones and concluded that digital creation seems to have a positive influence on students' motivation during the task.

3.1. Related Work

Our main contribution is that we formalize the general approach of aggregating concept map data mathematically and that we are linking it to the psychological foundations of both learning and concept mapping which is often neglected in prior research. The analysis methods that we are presenting have been adapted to work with concept landscapes, but are otherwise not new. Prior research that is related to our work is mostly coming from two different directions: First, the analysis of single concept maps with methods similar to ours and second, the analysis of more than a single map.

There is a large body of previous work concerned with the analysis of single concept maps. Valerio et al. (2008), Leake et al. (2005), and Leake et al. (2004) present graph-theoretic approaches to analyzing e.g. the importance of single concepts in a map. Koponen and Pehkonen (2010) investigate the organization of concept maps with the help of physical graph models. The Pathfinder algorithm that is used later on has been applied to single concept maps before (Taricani and Clariana, 2006). Grundspenkis and Strautmane (2010) are analyzing concept maps based on the occurrence of subgraph patterns. Villalon and Calvo (2008) present a framework for mining concept maps from text documents like students' essays.

Also, there is some work that considers more than a single concept map. For example, Larraza-Mendiluze and Garay-Vitoria (2013) investigate a set of concept maps with techniques of social network analysis by forming a weighted graph from a set of maps collected in a lecture. Yoo and Cho (2012) present the results of applying algorithms from subgraph mining to a set of concept maps. As part of the "Betty's Brain" software, the development of students' causal maps are visualized by aggregating the single maps (Biswas and Sulcer, 2010).

3.2. Possibilities and Limitations of Concept Maps

A concept map has two facets: First, it encompasses a *set of propositions*, i.e. the factual "content". Second, it encompasses the *structural configuration* as a network of linked concepts. The task of concept mapping has close resemblance to the mental function of "integration" which "refers to the process of finding a relationship that meaningfully links two concepts together"

(Solomon et al., 1999, p. 102). Therefore, it is important that the concept mapper is specifically asked to provide a linking phrase (even though this may be ignored in subsequent analysis) and that the task is open-ended (i.e. there is not just a list of possible linking phrases to choose from). The task of creating a concept map from scratch without any further constraints has been referred to as the "gold standard" of concept map assessments (cf. Yin et al., 2005). However, Cañas and Novak (2012) list several common restrictions, among which is providing a list of concepts that should be used for map construction. They note that "research has shown that the same students construct better maps when given a *list of concepts* [...]. More specifically, even if the number of concepts is similar in both cases (with the given list and without), the structure of the maps is different" (Cañas and Novak, 2012, p. 5).

For our work, we assume that the most gain from analyzing concept maps comes from the *structural information* of the maps (i.e. the interrelations between the concepts) and not from the list of propositions that they encode. Inferring domain knowledge from an assessment of the structural configuration of conceptual knowledge has been shown to be valid in many different areas (cf. Trumpower et al., 2010). Beyond the importance of structured knowledge that has been hinted at in the last section, it has also been explicitly noted in the literature that the structure of a concept map is an important aspect (cf. Cañas and Novak, 2012). McClure et al. (1999, p. 491) assert that it is "the organizational component captured by concept maps that may allow teachers to identify and correct student misconceptions." Koponen and Pehkonen (2010, p. 16709) have analyzed concept maps and conclude: '[W]ith more connections the structure also becomes more ordered. This suggests that students that are able to provide more connection (having more knowledge at their command) are also better at organizing that knowledge."

From the perspective of classical test theory, a concept map is a "noisy" measurement of a person's knowledge. That measurements of knowledge are inherently "noisy" has been observed at least for expert knowledge: "[A]lthough different experts may show variability in their judgments of concept relations, this variability often appears to be the result of random error rather than systematic differences in thinking" (Trumpower et al., 2010, p. 8).

One of the advantages that aggregating concept map data provides is that we expect an improvement in validity: Many variables influence the externalization, including the volition of the person, the prior experience with concept mapping, the time given to create the map and more. While the extent of "noise" in concept mapping, in general, is not known, it is at least plausible that for many of the influencing variables, the "noise" over many measurements is normally distributed around 0 and does therefore not systematically add up. This is especially true for personal variables that influence the externalization. For example, when a group of persons is creating concept maps, it is reasonable to assume that the motivation will vary from persons with only little motivation for the task to persons with rather high motivation. Overall, therefore, the influence of a lack of motivation for concept mapping that is detrimental when assessing only a single, unmotivated person is far less severe for a group where typically only a few persons will be unmotivated. The same can be expected to hold, within reasonable bounds, for the training necessary to create concept maps and the time given to create the map.

Also, due to the nature of concept mapping, it is in general not possible to deduct anything for missing elements in a concept map. If concepts or propositions are not contained in a concept map, the knowledge structure of the person might nevertheless contain them. When analyzing groups and concept landscapes, this is different: If a single person is missing a particular concept or connection, then this may be for a variety of reasons unrelated to the learning or the instruction this person received. If, however, every person of, for example, a class misses a particular concept or propositions, it becomes far more likely that this is due to their specific shared learning environment or input they received. Therefore, it is expected that the combination of many concept maps for analysis effectively reduces the influence of the main variables that may negatively impact the externalization.

3.3. Preliminaries

Following the basic definitions of graph theory, e.g. (Balakrishnan and Ranganathan, 2012), it is self-evident that concept maps in their most general form can be modeled mathematically as little more than a labeled, directed graph, as already noted in e.g. (Leake et al., 2005) or (Koponen and Pehkonen, 2010). The concepts form the nodes of the graph, and the labeled arrows form the edges. Care must be taken, however, that a concept should appear only once in a concept map. The rest of this article will be based on the following model: A concept map $CM = (V, E, L_v, L_e)$ consists of a finite (and usually non-empty) set of vertices and edges that form a directed, labeled graph with two labeling functions $L_v : V \to \Sigma^*$, $L_e : E \to \Sigma^*$ that map the set of vertices and edges to their labels taken from the set of all words Σ^* formed over some appropriate alphabet Σ. Additionally, it must hold that L_v is injective, i.e. that $L_v(x) = L_v(y)$ if and only if $x = y$. The labels can be defined more precisely (e.g. by formal languages), to restrict concepts to a single word, for example. Also, though not very common, self-loops are allowed in this definition and are sometimes found in real-world concept maps, e.g. when representing the statement "objects are communicating with objects" in the context of object-oriented programming. Given a set of concept labels S, it is sometimes convenient for analysis to restrict an existing map to the concepts of S. This is equivalent to using the subgraph induced by the vertex set $L_V^{-1}(S)$ for an appropriately defined inverse function of L_V.

In the analysis of concept maps, it is often helpful and sometimes required that the propositions of a map be scored. This score can then be used for a more detailed evaluation. A scored concept map $CM(V, E, w, L_V, L_E)$ is a weighted, directed, labeled graph, with a weight function $w : E \to \mathbb{R}$, where $w(i)$ denotes the score of the proposition formed by edge i and its incident nodes.

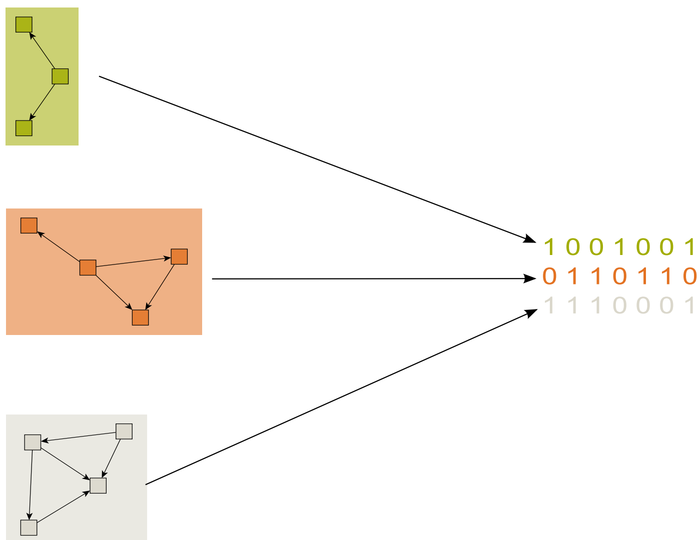
3.4. Definition

While in theory, any set of maps can be used to form a concept landscape, in practical settings the maps typically have been created by several persons in the same (educational) context, e.g. all students of a class at the same point in time. The aggregation itself can then happen in one of two ways:

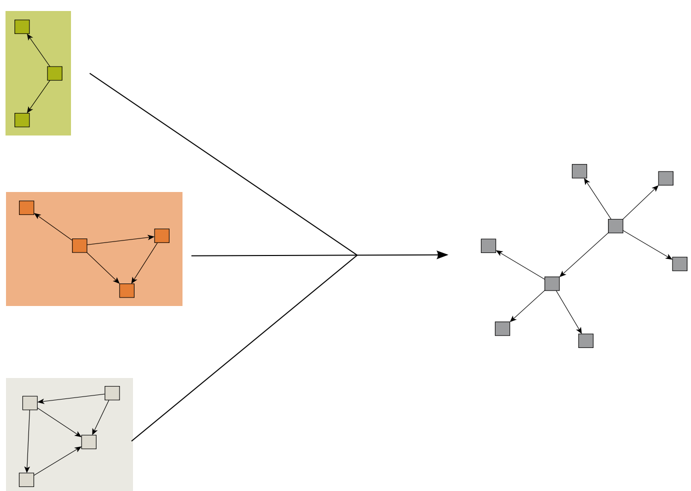
Accumulation The maps are still treated as individual entities; however, information contained in them is combined in some form for analysis. Often, each map will be transformed into a numeric vector, and a matrix will then be formed from all these vectors.

Amalgamation The individual maps are combined into a single graph that is then analyzed. Often, this newly formed graph will be weighted to reflect some of the statistical properties of the original set of concept maps.

Fig. 2 visualizes the difference. An accumulation still treats each concept map as an entity that is recognizable in the data. When amalgamating, the single maps are integrated into a whole, and the individual maps can no longer be identified.



(a) Accumulation



(b) Amalgamation

Figure 2: Accumulations still treat each map as an identifiable entity in the resulting data. Amalgamations result in a graph or concept map. In both cases, the newly formed data is the input for the subsequent analysis steps.

More formally, we assume a set consisting of k concept maps: $M = \{CM_1(V_1, E_1, w_1, L_V^1, L_E^1), CM_2(V_2, E_2, w_2, L_V^2, L_E^2), \ldots, CM_k(V_k, E_k, w_k, L_V^k, L_E^k)\}$. Typically, it makes sense to restrict the maps to a common set of concepts for analysis, but this is not necessary.

For accumulations, each constituent map is still an identifiable part of the "whole". In its most basic form, this means that a function is applied to each concept map which extracts any relevant features for future analysis, and the concept landscape is then simply the set of the function values, i.e. $CL = \{f(CM_1), f(CM_2), \ldots, f(CM_k)\}$ for some function $f : M \to X$ and an appropriately defined features space X. A convenient special case of this notion is to model the landscape as a (numerical) matrix with k rows, where row i is a vector that represents map CM_i. So, an accumulation is a matrix $CL \in \mathbb{R}^{k \times j}$, such that the i-th row is defined by the value $f(CM_i, \theta)$ of a mapping function $f : M \times \Theta \to \mathbb{R}^j$ defined appropriately, for some (optional) parameter space Θ.

The following mapping functions have been applied successfully in experimental studies:

Concept Vector Assuming that each of the k concept maps shares the same (ordered) set of concepts C, the concept vector encodes for each concept in C whether or not a concept map has an incident edge to this concept. Formally, the mapping function $f_C : M \times \{C\} \to \mathbb{B}^{|C|}$ maps a concept map on a binary vector such that:
$f_C(CM_i, C) = (v_1, v_2, \ldots, v_{|C|})$ with $v_k = \min(|\{(x, y) \in E_i : x = {L_V^i}^{-1}(c_k) \vee y = {L_V^i}^{-1}(c_k)\}|, 1)$. If the concept map is scored, it is also possible to only regard edges above a certain score limit, for example.

Edge Vector Again, assuming that each of the k concept maps share the same set of concepts C, the edge vector encodes the presence or absence of each of the possible edges that are appearing in a complete graph over the set of nodes C. Let E_C be the (ordered) set of edges of this complete graph and $e_k = (x_k, y_k)$ the k-th edge of this set, then the mapping function $f_E : M \times \{E\} \to \mathbb{B}^{|E_C|}$ is defined as $f_E(CM_i, E_C) = (v_1, v_2, \ldots, v_{|E_C|})$ with $v_k = 1$ if $({L_V^i}^{-1}(x_k), {L_V^i}^{-1}(y_k)) \in E_i$ and 0 otherwise. Again, edge scores can also be incorporated to filter out edges below a certain score, for example.

Graph Similarity A mapping that doesn't encode a particular feature of the concept map is built upon the function $f_S : M \times \{M\} \to \mathbb{R}^k$ which is defined as $f_S(CM_i, M) = (v_1, v_2, \ldots, v_k)$ such that $v_j = S(CM_i, CM_j)$ for some measure of graph similarity S. In other words, the vector encodes the similarity between CM_i and all maps of M.

In our case studies, we used a measure of graph similarity that has been reported by Goldsmith and Johnson (1990) to work especially well for investigating structural knowledge. They conclude that - in this context - the structural similarity can best be measured by comparing the neighborhoods of all nodes between two given graphs. Specifically, let $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ be two graphs that share a common set of vertices $V = \{v_1, v_2, \ldots, v_n\}$, Then the similarity of G_1 and G_2 can be calculated by summing the similarity of each neighborhood as follows:

$$C = \frac{1}{|V|} \sum_{v \in V} \frac{|N_{G_1}(v) \cap N_{G_2}(v)|}{|N_{G_1}(v) \cup N_{G_2}(v)|}$$

The value of C will vary between 0 and 1, where 1 denotes structural identity and 0 denotes a completely different structure. Note, that the fraction is undefined, if the union of the neighborhoods is empty, i.e. when both nodes are unconnected in both graphs. As they are then structurally identical, however, it is convenient to define the value of the fraction as 1 in this case. Other measures of graph similarity can be used just as well with our approach.

An amalgamation, by definition, is a graph that is formed in a precisely defined way out of a set of concept maps. There are several possible ways of forming such a graph. In its most abstract form, the set of nodes is the union over all nodes of the constituent concept maps $V = V_1 \cup V_2 \cup \ldots \cup V_k$. Sometimes it is convenient to first restrict all concept maps to a common set of concepts C (such that $V = V_1 = V_2 = \ldots = V_k = L_V^{1^{-1}}(C)$). The set of edges can be defined as the union as well: $E = \bigcup_{i=1}^{k} E_i$. A vertical amalgamation is simply the graph based on these sets of nodes and edges together with some weighing function $CL = (V, E, w)$. A simple weighing function is defined as: $w(e_i) = |\{1 \leq j \leq k | e_i \in E_j\}|$, for $w : E \to \mathbb{N}$. This counts the number of original maps in which certain edges are present. If a lower value should indicate a "more common" edge (as is the case for example in the Pathfinder algorithm used below), one can simply use the number of concept maps plus one and subtract from this the number of concept maps in which the edge is present: $w(e_i) = k + 1 - |\{1 \leq j \leq n | e_i \in E_j\}|$ for a combination of k concept maps. An edge that appears in every map is assigned a weight of 1 and an edge that appears in only 1 map is assigned a weight of k. Alternatively, instead of forming the simple sum, some additional filtering or transformation may be applied. For example, if the maps are scored, only edges with a score higher than some threshold could be used. Or only edges that are used in more than a given number of maps might be regarded in the analysis. A filtering that regards only edges with a score higher then t is defined as $w_s(e_i) = |\{1 \leq j \leq k | e_i \in E_j \wedge w_j(e_i) > t\}|$. Finally, a filtering that regards only edges that appear in more than t maps is defined as $w_t(e_i) = max\{|\{1 \leq j \leq k | e_i \in E_j\}| - t, 0\}$.

4. Working with Concept Landscapes

Working with concept landscapes means probing into knowledge structures. As knowledge isn't something that's easily quantifiable, every approach will usually rely on some form of qualitative evaluation in the end. Concept landscapes and related data mining methods are used to support or enable the qualitative, interpretive steps of a research process. The workflow in a typical research setting is:

1. Decide whether to use accumulation or amalgamation.
2. Decide on a specific transformation of the data.
3. Use the chosen method on the concept landscape.
4. Analyze and interpret the results - this may encompass repeating the steps for one or several subsets of the data that have been identified.

There are two aspects that can be of interest in concept landscapes:

1. The **inherent differences** in the knowledge structures. This amounts to finding subgroups of persons who differ in the way their knowledge structures are organized. This typically

cannot be done "manually" and requires data mining approaches. To identify inherent patterns in a set of concept maps, it is necessary that the single maps are still identifiable entities in the concept landscapes. Therefore, this works usually only with **accumulations**.

2. The **common elements** of the knowledge structures. That means identifying the *salient structural information* inherent in a concept landscape, e.g. to make a manual inspection of the results feasible. Again - this typically requires the support of data mining methods. To identify salient common elements, it is usually best to use **amalgamations**.

An analysis can (and often will) encompass several steps. For example, when applying cluster analysis to focus on inherent differences in knowledge structures, it makes sense to afterward look for common elements among the identified clusters to better characterize them. Our formal definition of the concept landscape framework encompasses typical research scenarios. As reported in the related work section, sometimes a weighted graph has been formed from concept maps. This corresponds to an amalgamation using any of the desired ways to define the edges. When investigating the development of concept maps along with a lecture, for example, this corresponds to an accumulation that does not result in a matrix but, for example, maps each concept map onto itself (e.g. for visualization purposes) or onto some feature like the number of concepts and edges.

In the following, we present two analysis methods - cluster analysis and the use of Pathfinder networks. These methods have been tried on real world data and found to yield interesting results. Each section will include a case study that shows how the methods can help in educational research.

The data of these case studies was collected at two different occasions:

1. Students entering Technical University of Munich with a major of CS. The students received a list of 40 core CS concepts that have been extracted semi-automatically based on their frequency of appearing in the CS curriculum of secondary schools in the German federal state of Bavaria. The concepts are typically also appearing in the lectures of the first semester of CS studies. Students were asked to draw a concept map as part of their schedule of their first day at university.

2. All students taking part in a voluntary pre-course before the start of the first semester for CS majors also at the Technical University of Munich. These students received a list of concepts extracted manually from the printed material of the pre-course.

Data was collected in different years, so there is no overlap between these groups. In both cases, the participants received a short introduction to concept mapping and were given an exemplary concept map with non-CS concepts as part of their assignment. The students had about 40 minutes time and were asked to create a concept map based only on those concepts of the given list that they were familiar with.

We do not expect students to have had a prior exposure to concept mapping in general - there may be singular cases though. Based on the quality of maps - that is also reflected in the results presented below - we do not expect a lack of experience in concept mapping to have a negative effect on the analysis, as concept mapping is a technique that is reasonably well grasped by an introduction and an example.

4.1. Pathfinder Networks

The most basic form of an amalgamated concept landscape is a weighted graph with the edges reflecting how "common" a connection between two concepts is in the chosen set of concept maps. This graph is - by itself - not suited for qualitative inspection as it is usually very dense. The structural information is encoded in the edge weights only.

It is possible to extract information from this graph by all kinds of graph theoretic approaches. For example, centrality measures like the "betweenness-centrality" of Freeman (1978) yield interesting information about the importance of single concepts for the organization of the knowledge structure. Another interesting option is to use community detection algorithms, e.g. the one described in (Clauset et al., 2004), which produce a partition of the concepts into subgroups based on their interconnections.

However, it is also worthwhile to bring out the most dominant structural connections of the complete landscape to interpret it - for example when comparing concept landscapes of different groups. This amounts to pruning the dense graph such that the salient structural connections remain and become visible. A simple method would be to filter out all edges with a weight lower (or higher) than a chosen threshold, for example by using the weighing function w_t defined above for some appropriate value of t. A more elaborate way would be to create a minimal spanning tree of the graph.

A *Pathfinder network* is a generalization of a minimal spanning tree and particularly suited for our context. Pathfinder networks are based on the psychological model of a "network." Schvaneveldt et al. (1989, p. 252) note that "networks entail the assumption that concepts and their relations can be represented by a structure consisting of nodes (concepts) and links (relations). Strength of relations are reflected by link weights and the intentional meaning of a concept is determined by its connections to other concepts." It is self-evident that concept maps and amalgamated concept landscapes fit that definition.

Algorithmic methods can then be used to analyze such a network, or extract salient structural features. The Pathfinder algorithm is one such method. An alternative is, for example, *multi-dimensional scaling* (MDS) (cf. Bartholomew et al., 2008). The lengths of paths in the Pathfinder network contain information about how "close" or similar the connected concepts are in the original data. From a graph-theoretical point of view, constructing a Pathfinder network is simply an algorithmic method of (edge-)pruning a graph by keeping all nodes and systematically removing edges.

The advantage of using the Pathfinder network for the analysis of concept landscapes is twofold: In contrast to other scaling techniques like MDS, it can work directly on a graph as input, and it also produces a graph, making it suitable for the format of concept landscapes. Also, its original intention is directly related to analyzing the structure of conceptual knowledge, in contrast to other graph pruning techniques, like minimal spanning trees or the simple edge removal based on some threshold as in (Larraza-Mendiluze and Garay-Vitoria, 2013).

The actual Pathfinder algorithm based on matrix operations is described by Dearholt and Schvaneveldt (1990). The Pathfinder network is a graph that consists of the same nodes and components as the input graph but uses only a subset of its edges, with their weights preserved. The edges are chosen such that an edge remains in the graph if and only if it does not violate the triangle inequality when considering at most q intermediate steps (the resulting graph is q-triangular) and when using the *Minkowski-* or *r-metric*) as measure of distance (cf. Dearholt and Schvaneveldt, 1990). The weight of a path consisting of edges $e_1, e_2, \cdots, e_k$ with weights

$w_1, w_2, \cdots, w_k$ according to the r-metric ($r > 0$) is defined as: $(w_1^r + w_2^r + ... + w_k^r)^{1/r}$. For $r = 1$ the r-metric defaults to the sum of the single edge weights, for $r = 2$ it is the Euclidean distance and for $r = \infty$ the path weight is the maximum edge weight along the path (cf. Dearholt and Schvaneveldt, 1990). An appropriate definition of the edge weights must somehow reflect an edge's "commonness" with a lower weight defining a higher commonness. One possibility is the weight function presented above that uses "$k + 1-$ number of maps with a given edge" as weights.

The Pathfinder network is dependent on the choice of the parameters q and r. In general, higher values for either parameter will produce sparser networks. From a perspective of graph pruning, the typical choice for the value of q is $|V| - 1$. Using smaller values would be useful only if there were a theory about why the structural information of paths that are exceeding a certain number of intermediate steps is not as important as the information of shorter paths. For r, the extreme values of 1 and ∞ are most appropriate. Euclidean distance ($r = 2$) might be interesting if it can be assumed that the spatial placement of concepts in a map is an important aspect of the organization of knowledge.

The pruning of the Pathfinder algorithm differs fundamentally from a manual removal of edges with a low (or high) weight: A Pathfinder network will always contain the same components and the same nodes as the original graph. This has two consequences for analysis:

First, it is paramount to keep in mind that it is in general not possible to compare, for example, two different pairs of concepts based on the fact that the edge between one of these pairs was removed by the Pathfinder algorithm, but not between the other. If a concept has only one incident edge, it may happen that every edge that is pruned by the Pathfinder algorithm had a lower weight than this one. In the most extreme case, a concept that is used in e.g. only one map of a concept landscape will remain in the Pathfinder network using the one edge that is present in the data, because pruning this edge would create a new component. However, this can hardly be taken as an indicator that this edge is representative of the "common" knowledge structure, as it is only there as an artifact of an idiosyncratic knowledge construct that is not removed by the analysis method.

Therefore, second, it is important to filter out these idiosyncrasies by manual removal. This can be done with a weighing function similar to w_t - by setting the weights of filtered edges to ∞. Alternatively, this can be done directly on the weight matrix before applying the Pathfinder algorithm. For example, rows with a sum above a given threshold can be removed (i.e. concepts removed), or entries that have too high a value can be set to ∞ (i.e. edges removed). There is no best way to arrive at a threshold for filtering, but using percentiles provide a reasonable approach. It seems advisable to try several threshold values and inspect the effects on the network.

4.1.1. Case Study 1

The research presented here is part of a larger research project that was concerned with the impact of compulsory computer science (CS) education in secondary school on the structural knowledge of beginning students. Such information is valuable for the design of pre-courses or for the lecturers in the first semester as they can then accommodate the prior knowledge of their students more effectively. Prior knowledge is a highly important aspect of successful learning (cf. Ausubel; Gurlitt and Renkl, 2000; 2010).

Most of the CS students entering our university have attended secondary school in our federal state. In this state, a compulsory subject of "Informatics" had been introduced some years ago - together with a reform of the school system itself which led to a reduction from 9 to 8 years of secondary education. Therefore, at one year, the last class of the old system (9 years, no compulsory CS education) and the first class of the new system (8 years, compulsory CS education of - typically - 4 years) enrolled at university in the same term. We chose this opportunity to ask all beginning CS students at our university to draw a concept map to investigate their understanding of basic CS concepts.

In total, 590 concept maps were collected, and after excluding empty and nonsensical maps, roughly 350 remained for analysis.

Based on the prior formal CS education, the data naturally splits into two groups. We used those groups to form a concept landscape from each. The maps were all restricted to the 40 concepts that were given to the students. Then, the maps of each group were amalgamated canonically - the edge weights reflect the number of maps that show a connection between the two concepts.

The Pathfinder algorithm was applied to each concept landscape. The parameters were set to $q = 39$ and $r = \infty$ to create the sparsest possible network. Also, as described above, a manual filtering step should occur before applying Pathfinder to filter out idiosyncrasies. In our case, all edges that appeared in less than 10% of the maps (of each group) were removed.

Figure 3 and 4 show the resulting graphs (unconnected concepts have been removed).

These two networks allow a qualitative interpretation of the structural differences as they are easily grasped visually. The colored edges show areas of interest:

Data structures and abstraction (red and green) While the green edges show that both groups have a similar grasp of the concept of *data structure* and the examples that were included in the list of concepts (*graph, tree, array, list*), there is a clear difference concerning the connection of this "block" to the rest of the knowledge structures. The students who had a prior CS education in school are showing a clear distinction between the theoretical notion of *data structures*, the concept of *data*, and the more technical concept *working memory* (where data often resides) - as seen in the red edges. The students without prior, formal CS education do not show this distinction, and the red path is "collapsed" in their knowledge structure.

Approach to programming (turquoise) In the subject "Informatics", the students are learning to program in a strictly object-oriented fashion. This is visible in their knowledge structures in the turquoise edges as the two programming constructs *loop* and *conditional statement* are connected to *method*. The students have practically only been programming methods that reside in *classes* which in turn are the basis of a *program* in an object-oriented programming language. The other group connects these concepts via *variable* directly to *program* which is more akin to a procedural programming language that they might have taught themselves (most beginning CS students have at least some prior programming experience). They also do "know" about the object-oriented concepts, but it is less pivotal for actual programming to them.

Formal languages (blue) A typical set of core CS concepts deals with the difference between *syntax* and *semantics* and the formal definition of (*programming languages* with *grammars*. The students with prior education show that these concepts are integrated into their knowledge structure about CS. The other group of students is missing *semantics* completely, and *grammar* has no connection besides *programming language*.

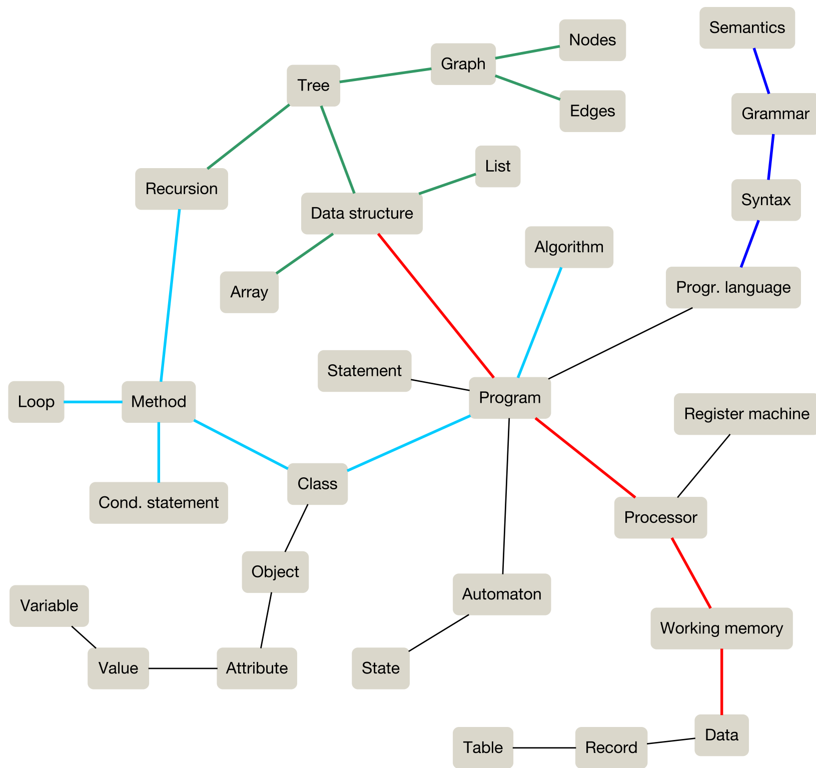


Figure 3: The Pathfinder network of the students with compulsory CS education in school.

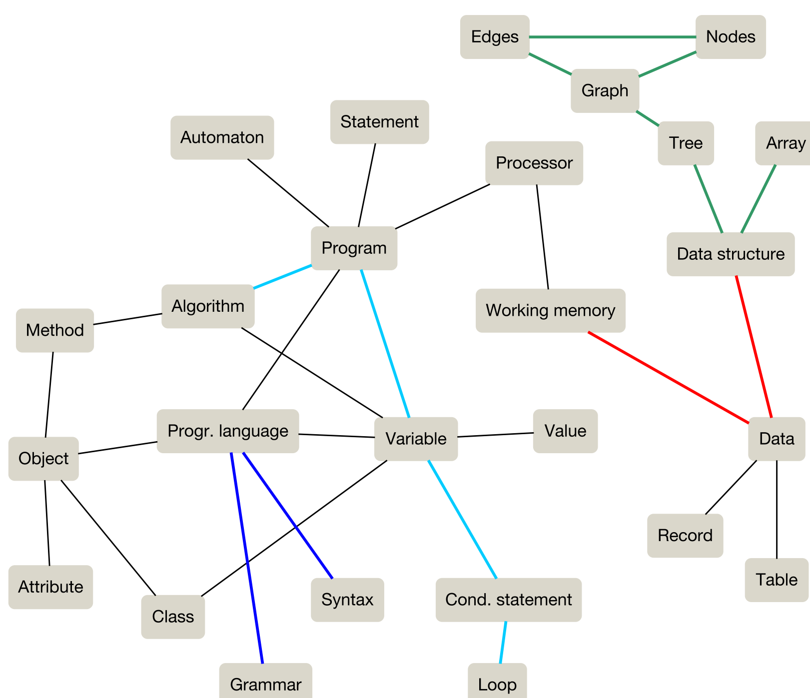


Figure 4: The Pathfinder network of the students without CS education in school.

There are also similarities like the obvious importance of the concept *program*. When measuring the betweenness-centrality for the original concept landscapes (before applying the Pathfinder algorithm) *program* receives the highest value for both groups. However, the second highest value already differs: The group with prior education has the concept of *class* taking this place while the other group has *processor* - again indicating the results of the clear object orientated approach in the school subject and the more technical approach to CS from the other students.

4.2. Cluster Analysis

In the prior case study, a partitioning of the data due to a personal variable of the participants (their prior formal CS education) was obvious. However, in other cases, the pattern that is inherent in the data is not as readily observable. Cluster analysis is a typical research method in exploratory analyses, as it identifies patterns in the data which can then be interpreted (cf. Bartholomew et al., 2008). There is a plethora of clustering algorithms and approaches available, and many of them will work with concept landscapes.

One large group of clustering algorithms is similarity based. These algorithms typically take a distance matrix as an input and produce a clustering based on these distances. An often-used algorithm that is appropriate for our kind of input data is *k-medoids* as described by Kaufman and Rousseeuw (2005). To arrive at a distance matrix that encodes the pair-wise similarity of observations (i.e. concept maps) we can use the graph similarity based mapping function defined above. The resulting numerical matrix already is a distance matrix. When using mapping functions that encode the presence or absence of features, the resulting binary matrix is not yet a distance matrix. It can be transformed into one by applying a norm on the difference between each pair of rows. When using the concept vectors to form the binary matrix for example, the Manhattan distance (L_1 norm), usually defined as $\sum_{i=1}^{n} |x_i|$ for an n-dimensional vector x effectively encodes the number of concepts in which the two concept maps differ structurally - in the sense that only one of the two maps has an incident edge to a concept. A higher value corresponds to greater structural differences, making the maps more “dissimilar”.

Similarity-based clustering algorithms usually get the number of clusters they should produce as a parameter. Therefore, to determine the optimal number of clusters, some form of quality measure must be applied. Typical choices in literature are the *Calinski-Harabasz pseudo F-statistic* (cf. Gordon, 1999) or the Gap statistics (Tibshirani et al., 2001). Depending on the method chosen it is also necessary to include the special case of only one cluster, i.e. the case that the data is not uniformly enough to warrant a clustering. This can be checked, for example, by the Hopkins index (Han and Kamber, 2010), or by the Duda-Hart test (Duda and Hart, 1976).

A model-based approach allows identifying the optimal number of clusters, including the case of just one cluster, based on the likelihood of a set of parameters for an explicit stochastic model of the data. Clustering is then no longer the task of assigning observations to clusters. Instead, the task is to identify model parameters such that the likelihood of the model is maximized (cf. Han and Kamber, 2010). To avoid overfitting, evaluating AIC (Akaike, 1974) or BIC (Schwarz, 1978) instead of the likelihood of the model is a common approach.

A stochastic model that can be used for describing a binary matrix that results from encoding the presence or absence of features is a multivariate Bernoulli-mixture model (MBMM). This allows for correlations among the entries of different columns. These exist in our case by definition - for example for the concept vector, a 1 in one column will always indicate the presence of another 1, as each edge has two incident concepts. Following e.g. Stibor (2008), such higher order correlations can be modeled by using multivariate Bernoulli mixtures. The parameters of a MBMM consist of m vectors $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_m)$ - with each vector $\Theta_i = (p_1^i, p_2^i, \ldots, p_k^i)$ describing a single multivariate distribution of k different Bernoulli distributions with probability p_j^i - and m mixing coefficients $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$. It holds that $\sum_{i=1}^{m} \alpha_i = 1$. The α_i encode the contribution of each distribution to the observed result. Each Θ_i consists of k values, so an MBMM has $m \cdot k + m - 1$ parameters as there are $m - 1$ mixing coefficients that are free to choose. These must be chosen (optimally) to identify a clustering. The probability function for an observation $x = (x_1, x_2, \ldots, x_k) \in \{0, 1\}^k$ is $P(x) = P(x|\Theta, \alpha) = \sum_{i=1}^{m} \alpha_m P(x|\Theta_m)$, where the probability of the multivariate distribution with parameters Θ_j, $P_j(x) = \prod_{i=1}^{k} (p_i^j)^{x_i} (1 - p_i^j)^{1-x_i}$.

Finding the optimal parameter values for given input data corresponds to maximizing the likelihood of the observed data. A typical approach is to use the EM algorithm (Dempster et al., 1977). The E- and M-steps in our specific case are (see e.g. (Wolfe, 1970) or (Stibor, 2008)):

E-Step Calculate for each mixture component m and each observation of the data x_i the posterior probability using Bayes theorem:

$$P(m|x_i, \Theta, \alpha) = \frac{P(x_i|m, \Theta, \alpha) P(m)}{P(x_i)}$$

M-Step Calculate for each mixture component m new parameter values Θ'_m, α'_m using the current values Θ_m and α_m, by:

$$\alpha'_m = \frac{1}{n} \sum_{i=1}^{n} P(m|x_i, \Theta, \alpha)$$

and

$$\Theta'_m = \frac{1}{n\alpha'_m} \sum_{i=1}^{n} P(m|x_i, \Theta, \alpha) x_i$$

The computation can be effectively implemented using matrix operations. The output of the EM-algorithm consists of the locally optimal (i.e. most probable) values of the probabilities Θ and mixing coefficients α. These values can be used to calculate the posterior probability (as given above) for each observation and each cluster. This, in other words, is the probability for each observation to belong to any of the clusters.

When clustering the data of concept landscapes, it is always important to analyze the clusters afterward as the clustering algorithms itself may or may not produce clusters that are worthwhile from a researcher's perspective. Also, it is important to note that in general the clusters can only be interpreted by additional analysis steps. Otherwise it may happen that the algorithm produces a clustering based on the influencing factors of the externalization that the concept landscape tries to minimize. If, for example, the maps of one cluster are very sparse, then maybe this cluster is formed by all the persons who didn't understand concept mapping or weren't motivated at the time of creation to produce larger maps.

The next two subsections are presenting two case studies - the first uses k-medoids clustering, the second uses MBMM clustering.

4.2.1. Case Study 2

We use the same data as in the first case study. However, instead of manually forming clusters, we are now interested in differences in the structural knowledge that may not be due to students' prior education. To find out whether a clustering is worthwhile, we use both the Hopkins index and the Duda-Hart test. Both indicate that a solution of only one cluster is not optimal. The Calinski-Harabasz statistics points to an optimum of two clusters within the range of two to ten clusters. The Gap statistics indicates an optimal number of three clusters. So, the data clearly shows a tendency to form a small number of clusters at least for the chosen similarity measure.

We checked both the two-cluster solution and the three-cluster solution. The clustering is indeed picking up other traits of the knowledge structures, as the manual groups formed in case study 1 are not reproduced here. Instead, the students with and without prior formal CS education are split almost evenly over the clusters. A χ^2-test reveals no significant deviation from a uniform distribution in both cases. The clusters are of roughly equal size in both cases: 169 and 164 for two clusters and 116, 95, and 122 for three clusters.

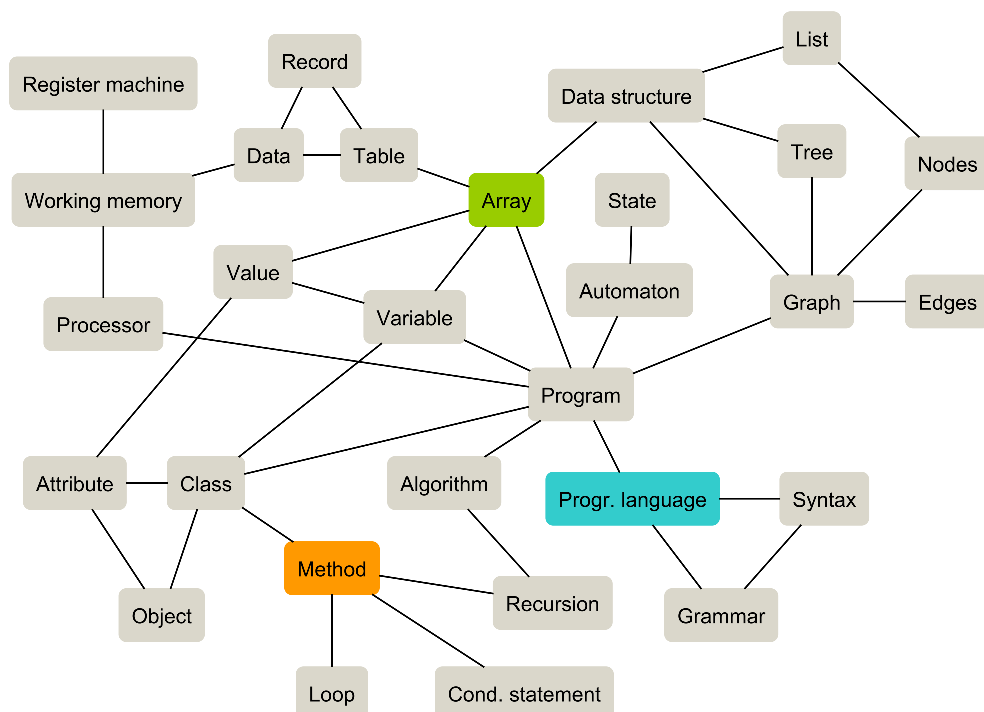


Figure 5: First cluster for three-cluster solution.

To investigate what separates the clusters, the maps of each cluster were amalgamated separately and the Pathfinder network using the parameter values that yield the sparsest graph were created from each landscape. Again, all edges that appeared in less than 10% of the maps of each cluster were filtered out beforehand and concepts that were left unconnected afterward were removed. The resulting Pathfinder networks of the three-cluster solution (Fig. 5, 6, and 7) are shown here. The concepts discussed below have been colored in each network.

As the distance matrix is based upon a measure of graph similarity that is defined by the

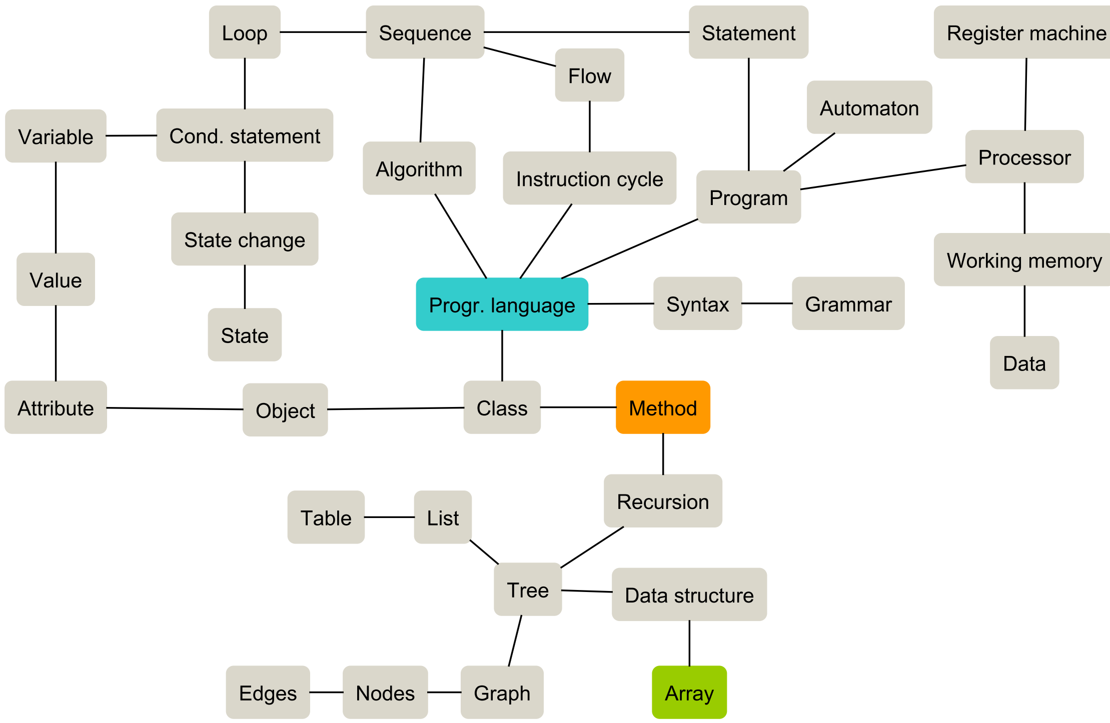


Figure 6: Second cluster for three-cluster solution.

similarity of neighborhoods, it makes sense to look how these neighborhoods differ between the clusters over the given concepts. The following observations can be made:

Programming Language Here, there is a distinct difference between the two- and three-cluster solution. While the concept *programming language* is always connected to *program*, it is typically connected to only *syntax*, *grammar*, or *semantics* in differing configurations. However, one of the clusters in the three-cluster solution has programming language additionally connected to *instruction cycle*, *algorithm*, and *class*. For this cluster, the concept programming language is overall placed very centrally in the knowledge structure.

Method As before, the clusters show some differences regarding the object-oriented concepts and their connection to other concepts. For the two-cluster solution, the concept *method* is only connected to *class* for one cluster but connect to *class*, *loop*, *algorithm* and *cond. statement* for the other cluster. The same trend is observable in the three-cluster solution: One cluster shows only a connection of *method* to *class*, one cluster has an additional connection to *recursion* and the third cluster has *class*, *loop*, *cond. statement* and *recursion* connected to *method*. So, the three-cluster solution reproduces the trend of the two-cluster solution but has an additional "intermediate" case.

Array For the concept *array* there is a similar distinction between the two- and three-cluster solution as for *programming language*: In the two-cluster solution, this concept is only connected to data structure for both clusters. The three-cluster solution however identifies a subgroup that treats this concept differently. One cluster shows the same trend, one

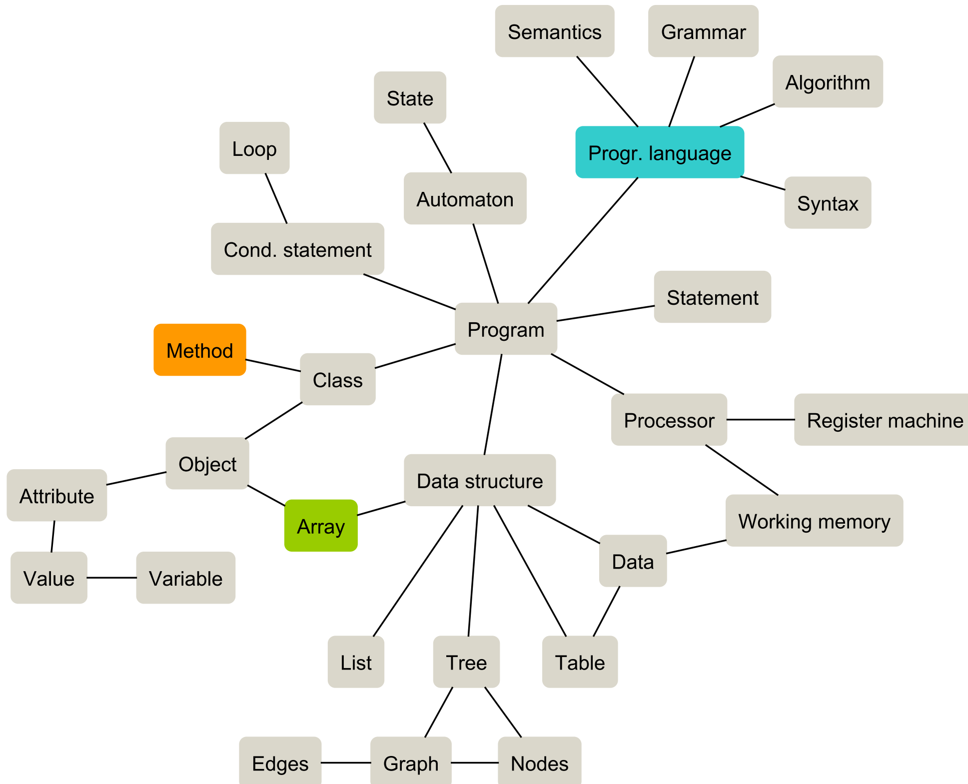


Figure 7: Third cluster for three-cluster solution.

additionally shows a connection to *object* and the third has additional connections to *value*, *variable*, *program*, and *table*

Overall, the three-cluster solution shows some interesting features. Each of the three clusters has a distinct separation in one region of the landscape: One cluster has a unique way of integrating *programming language* with the other concepts, another cluster does have a more elaborate way of connecting *array* with the other concepts. The third cluster has data structure connected with program - a unique feature of this group that also leads to a very thematically structured network, overall.

For the two-cluster solution, the difference between the clusters is mostly hinged on the different integration of *method* - similar to the distinctions found in the first case study. Also, both clusters are characterized by the appearance of concepts - or rather the remaining of these in the Pathfinder network - exclusive to one of the two clusters: *state* and *semantics* for the first cluster, *recursion* and *flow* for the second.

4.2.2. Case Study 3

A pre-course for beginning CS students - as suggested above - has been implemented in our university. It deals with the basics of object-oriented programming. Only minimal teaching input was given, and self-guided learning-by-doing was stressed. We extracted a list of concepts from the worksheets that the students used and asked them to draw a concept map based on this list after the course. In total, there were 75 maps that we could use for analysis. In contrast to the prior studies, for this case study, we scored each proposition manually with a score between 0 and 2 regarding their correctness. Using three values has proven to be a fast and reliable method of scoring. We only used edges with a score of 2 for the rest of the analysis.

We were interested in finding out how the knowledge structures of the students develop during their active experimentation with the concepts in the course. The resulting maps were accumulated, and the MBMM based clustering described above was used together with the mapping function that uses concept vectors. The AIC values indicate an optimum at two clusters. Both clusters are of equal size with 37 and 38 members respectively. In the next step, the maps of each cluster were accumulated separately, again using the mapping based on the concept vector. Then, for each row of the resulting matrices, the relative frequency of entries with a 1 (i.e. the relative frequency of concept occurrence) is plotted for both clusters, as shown in Fig. 8.

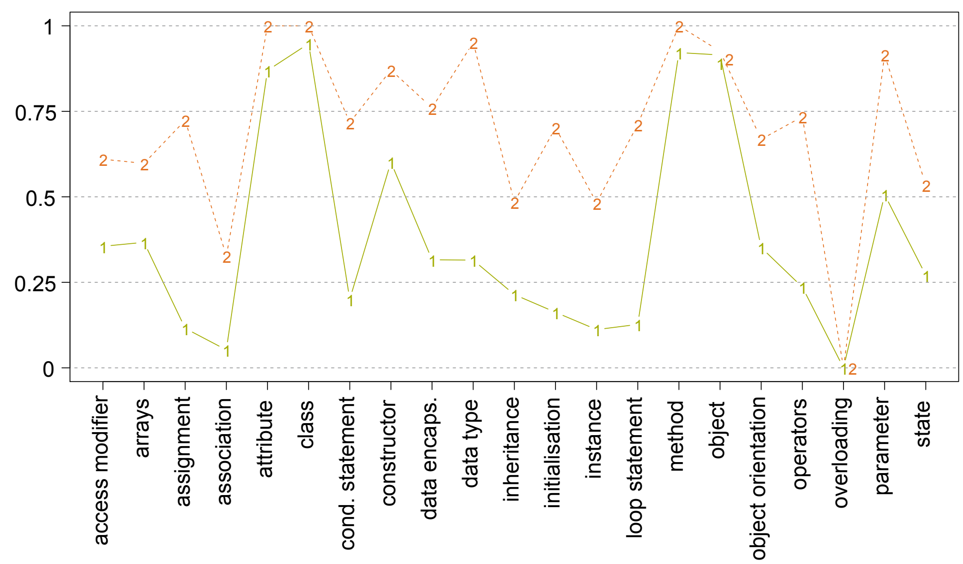


Figure 8: The probabilities of concept occurrence as identified by the MBMM clustering algorithm shown for both clusters.

The concepts with an inter-cluster difference of more than 0.5 of the corresponding probabilities are (in descending order of the difference): *data type*, *loop statement*, *assignment*, *conditional statement*, and *initialization*. This list is interesting insofar as there are no concepts present that are "purely" object-oriented (like *class*). Instead, the list contains the very basic concepts of procedural programming, namely assignment as well as the two control structures conditional- and loop statement. Also, the members of the first cluster tend to focus mostly on the basic concepts of object orientation in their maps.

From the perspective of the course designer, the results are encouraging as both clusters have incorporated basics of object-orientation in their mental models. The rest is just about how much they could take out of the course besides the basic concepts. For more details about this case study and further results of the analysis, see (Berges et al., 2012).

5. Discussion of the Case Studies

Since the first two case studies are based on the same dataset and different methods have been used in the case studies, it seems worthwhile to investigate how the different approaches to analyzing the data compare to each other[1]. Both the Pathfinder algorithm and a non-model-based clustering are by itself agnostic to the underlying data. The Pathfinder algorithm will always produce an output, as will a k-medoids clustering for a given k. The resulting graph(s) can then be interpreted, and if one is set one finding artifacts in the results, this will usually be possible. So, care must be taken to look at the data from as many angles as possible and look for those artifacts that remain stable throughout the differing approaches, e.g. by using a different number of clusters, a different algorithm, or a different measure of similarity.

While it is possible that a cluster analysis reproduces the "manual" clustering of case study 1, this would require the similarity measure to pick up on the specific traits that separate the two groups. As case study 2 shows, this is not the case for the graph similarity measure used here. This does not detract from the usefulness of the methods, however as they are exploratory in nature. It is obvious that there are several distinct groupings of the constituent maps that may be present in a given data set and that all reveal some interesting information. In the data of the first two case studies, one constant concerning the structural configuration has been the approach to object-orientation versus procedural programming, as exemplified by the concept *method* and its connections, as well as a tendency to place program as a highly central concept. Indeed, one of the major differences between the two-cluster solution and the three-cluster solution in case study 2 is the appearance of a cluster that does not place program as centrally as the other clusters. If configurations remain somewhat similar over different approaches and clusterings, then these can indeed be taken as features of interest in the concept landscape that are not just the product of the particular process that has been employed to obtain the results.

One can also wonder which clustering method is appropriate in each case. The results will differ as MBMM uses a different approach based on different features than k-medoids. Again, it is probably best to try both approaches and look for similarities. It is also possible, however, to argue for one or the other method based on the hypothesis that one has. The MBMM approach seems reasonable in case study 3, as we expected students to use only those concepts that they learned something about during the course. So, a clustering based on the *appearance* of a concept seems more appropriate than one that places more emphasis on the structural configuration. Using the same reasoning, the k-medoids approach seems better suited to case study 2 where we were expecting students to have a grasp of several of the given concepts already and we were more interested in how their knowledge structure looks like at the beginning of their studies.

6. Limitations and Future Work

The limitations of our current approach are mostly arising from two choices that we made: First, the choice of concept maps as a method of externalization and second, the choice to disregard the actual propositions and only focus on the structure.

Concerning our original research question regarding the knowledge structures of groups of learners, it seems that concept maps are a reasonable choice with promising results. However, alternative methods of externalization may also present viable alternatives for the specific task

[1] In particular, one reviewer of the original manuscript raised the question whether or not the results that are obtained with these methods can always be interpreted when looking at it "from the right angle".

of research on knowledge structures. As Cañas and Novak (2012, p. 249) put it: “[T]he educator needs to understand that the type and quality of the concept map may be more a reflection of the process and conditions under which the concept map was constructed than of the student’s understanding of the domain.” The structural information can also be externalized by relatedness judgments - for example - and the propositional information may be collected afterward as suggested by Albert and Steiner (2005). Madhyastha and Hunt (2009) analyze the similarity of concepts by using multiple choice data as the input.

Regarding the epistemological limitations, when concept mapping is mostly seen as a way of organizing and displaying knowledge, there are also better alternatives, due to the restrictions of concept maps. However, concept mapping has the inherent advantage that it is easy to learn for most people - in contrast to e.g. describing a knowledge base as an ontology - and it offers a direct benefit for the mapping persons - in contrast to e.g. relatedness judgments, especially when the mappers can keep or continue working on their concept maps. Also, concept maps are a valuable tool for formative assessment, in contrast to many other methods of externalizing knowledge. This makes concept maps valuable for applications in educational settings as there is more to be gained than “just” the externalized knowledge. On the downside, however, the method of concept mapping has to be explained to participants and may even require some form of training until they become fluent enough in the process. Also, the relative freedom of expression, especially for pen and paper based maps, often leads to highly variable results concerning the actual syntax and semantics of concept maps.

So far, we have only considered the structural configuration of the concept maps without regarding the actual propositions. As has been noted in the theoretical background, there is good reason to investigate this structure from a psychological point of view. Nevertheless, we are currently disregarding much information. A first step to incorporate the information has been hinted at in the last case study: By scoring propositions, we are at least evaluating some information of the proposition. Subsequent steps would most probably have to rely on some form of automatic text analysis since a manual inspection of each proposition is inconsistent with the idea of using data mining to process large bodies of data.

Concept landscapes should provide a framework for further research based on aggregated concept maps. Our definition of concept landscape as either accumulation or amalgamation encompasses every possible formalization. The specific definitions that were given here are, however, only one possible way of defining concept landscapes. Nevertheless, we assume that they are encompassing most, if not all, approaches towards analyzing concept landscapes with further methods - in particular with more advanced techniques from the EDM community. For example, there are many more features that could be encoded as the basis for clustering. Also, the analysis methods that were presented here are only a first selection of suitable methods. There are other ways of pruning graphs, other similarity measures, or ways of arriving at a distance matrix for clustering and other ways of defining latent class models for concept map data. Other methods can and will be investigated and used in the future. It is also important to acknowledge that concept landscapes are a tool best used in exploratory research settings. The results that are obtained can then serve as input for a more confirmatory approach.

7. Conclusion

This article has presented the notion of concept landscapes both in theory and practice. The goal was to present a working framework and analysis methods that allow investigating shared knowledge structures and that are based on a thorough theoretical basis. Concept maps have been chosen as the method for externalization because of their established validity and reliability, the existing psychological and educational theory behind it, and the additional benefit they can offer to participants (e.g. as learning aids). Based on these foundations, a novel view on the application of concept mapping in investigating the state and development of structural knowledge has been presented. Instead of focusing solely on the measurement of a single person, or many measurements of single persons analyzed in isolation, the data of a group of persons is aggregated. While motivational factors, the fluency regarding concept mapping itself, the specific location and time of the concept mapping task and other variables will, whether detectable or not, influence the results of each concept map, many of these influences can be expected to cancel out by aggregation.

The method of aggregation, as shown in Fig. 2, can either be a "loose" accumulation of maps that remain identifiable in the aggregated data, or it can be a more transformative amalgamation that results in a new graph that is influenced by the constituent maps of the landscape. Forming either an amalgamation or an accumulation does cover all possibilities as the concept maps either remain identifiable as entities in the aggregation or not.

Concept landscapes are only an abstract notion of a set of data items. In practical settings, it is the analysis and the actual method of aggregation that determines the usefulness of the approach. This article presented two basic analysis methods. Cluster analysis is best suited to identify inherent differences within an accumulation. While many clustering algorithms and similarity measures can be used, a latent model-based approach based on multivariate Bernoulli mixture models and partitioning methods using, for example, graph similarities as a measure of distance have been presented and shown to work well on actual data. Creating Pathfinder networks from concept landscapes is a way of pruning edges that allows discovery of salient structural information in amalgamations to identify the common structural elements in the data. The two central methods of clustering and Pathfinder networks can achieve useful results, but they will usually be accompanied by further ways of analyzing the results.

Typically, neither a purely quantitative nor a purely qualitative analysis yields optimal results in exploratory, fundamental educational research. Instead, both approaches have been used simultaneously and - one might say - non-dogmatically in order to improve on each other. For instance, finding clusters and then using a t-test to identify that one cluster possesses denser maps is a quantitative approach. Then continuing by searching graph communities in these clusters and inspecting them regarding their intersections and differences, based on the given subject matter context, is a qualitative analysis of the data. By using the strengths of both methods, a much broader variety of insights into the data of concept landscapes can be gained.

Having a working software tool-chain available opens the door for an economical and flexible way of monitoring certain aspects of educational processes in exploratory settings with minimal prior work. This includes a software based drawing of concepts maps, the analysis of the results and - possibly - the analysis of the learning stimulus. The R package CoMaTo offers support for the analysis step. Educators or researchers who are interested in the development of students' knowledge can incorporate concept mapping into their teaching activities, collect, and analyze the maps electronically with relatively minimal additional overhead.

REFERENCES

AKAIKE, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic*

Control 19, 6, 716–723.

ALBERT, D. AND STEINER, C. 2005. Empirical validation of concept maps: Preliminary methodological considerations. In *Proceedings of the Fifth International Conference on Advanced Learning Technologies, Kaohsiung, Taiwan, 5-8 July 2005*, P. Goodyear, D. G. Sampson, D. J.-T. Yang, Kinshuk, T. Okamoto, R. Hartley, and N.-S. Chen, Eds. IEEE Computer Society Press, Los Alamitos, California, 952–953.

AUSUBEL, D. P. 2000. *The Acquisition and Retention of Knowledge: A Cognitive View*. Kluwer Academic Publishers, Dordrect, Boston.

BALAKRISHNAN, R. AND RANGANATHAN, K. 2012. *A Textbook of Graph Theory*, 2 ed. Universitext. Springer, New York.

BARTHOLOMEW, D. J., STEELE, F., MOUSTAKI, I., AND GALBRAITH, J. I. 2008. *Analysis of Multivariate Social Science Data*, 2nd ed ed. Chapman & Hall/CRC and CRC Press, Boca Raton.

BERGES, M., MÜHLING, A. M., AND HUBWIESER, P. 2012. The gap between knowledge and ability. In *Proceedings of the 12th Koli Calling International Conference on Computing Education Research, Koli, Finnland, 17-15 November 2012*, M.-J. Laakso, Ed. ACM, New York, 126–134.

BISWAS, G. AND SULCER, B. 2010. Visual exploratory data analysis methods to characterize student progress in intelligent learning environments. In *Technology for Education (T4E), 2010 International Conference on*. 114–121.

CAÑAS, A. J., CARFF, R., HILL, G., CARVALHO, M., ARGUEDAS, M., ESKRIDGE, T., LOTT, J., AND CARVAJAL, R. 2005. Concept maps: Integrating knowledge and information visualization. In *Knowledge and Information Visualization*, S.-O. Tergan and T. Keller, Eds. Lecture notes in computer science, vol. 3426. Springer, Berlin and Heidelberg, 205–219.

CAÑAS, A. J. AND NOVAK, J. D. 2006. Re-examining the foundations for effective use of concept maps. In *Concept Maps: Theory, Methodology, Technology*, A. J. Cañas and J. D. Novak, Eds. Vol. 1. Universidad de Costa Rica, San José and Costa Rica, 494–502.

CAÑAS, A. J. AND NOVAK, J. D. 2012. Freedom vs. restriction of content and structure during concept mapping - possibilities and limitations for construction and assessment. In *Concept Maps: Theory, Methodology, Technology*, A. J. Cañas, J. D. Novak, and J. Vanhear, Eds. Vol. 2. 247–257.

CLAUSET, A., NEWMAN, M. E. J., AND MOORE, C. 2004. Finding community structure in very large networks. *Physical Review E 70,* 6, 066111.

COOKE, N. J. 1994. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies 41,* 6, 801–849.

DALEY, B. J., CONCEICAO, S. C. O., MINA, L., ALTMAN, B. A., BALDOR, M., AND BROWN, J. 2010. Integrative literature review: Concept mapping: A strategy to support the development of practice, research, and theory within human resource development. *Human Resource Development Review 9,* 4, 357–384.

DE JONG, T. AND FERGUSON-HESSLER, M. G. 1996. Types and qualities of knowledge. *Educational Psychologist 31,* 2, 105–113.

DEARHOLT, D. W. AND SCHVANEVELDT, R. W. 1990. Properties of pathfinder netowrks. In *Pathfinder Associative Networks*, R. W. Schvaneveldt, Ed. Ablex Pub. Corp., Norwood and N.J, 1–30.

DEMPSTER, A. P., LAIRD, N. M., AND RUBIN, D. B. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 39,* 1, 1–38.

DERBENTSEVA, N., SAFAYENI, F., AND CAÑAS, A. J. 2007. Concept maps: Experiments on dynamic thinking. *Journal of Research in Science Teaching 44,* 3, 448–465.

DUDA, R. O. AND HART, P. E. 1976. *Pattern Classification and Scene Analysis*, [28. Dr.] ed. A Wiley-Interscience publication. Wiley, New York.

FREEMAN, L. C. 1978. Centrality in social networks conceptual clarification. *Social Networks 1,* 1, 215–239.

GLASER, R. AND BASSOK, M. 1989. Learning theory and the study of instruction. *Annual Review of Psychology 40*, 631–666.

GOLDSMITH, T. E. AND JOHNSON, P. J. 1990. A structural assessment of classroom learning. In *Pathfinder Associative Networks*, R. W. Schvaneveldt, Ed. Ablex Pub. Corp., Norwood and N.J, 240–254.

GOLDSTONE, R. L. AND KERSTEN, A. 2003. Concepts and categorization. In *Handbook of Psychology*, I. B. Weiner, Ed. Vol. 4. John Wiley, Hoboken, 599–621.

GORDON, A. D. 1999. *Classification*, 2 ed. Monographs on statistics and applied probability, vol. 82. Chapman & Hall/CRC, Boca Raton.

GOULI, E. 2007. Concept mapping in didactics of informatics. assessment as a tool for learning in web-based and adaptive educational environments. Ph.D. thesis, National and Kapodistrian University of Athens, Athen.

GRUNDSPENKIS, J. AND STRAUTMANE, M. 2010. Usage of graph patterns for knowledge assessment based on concept maps. *Scientific Journal of Riga Technical University. Computer Sciences 38,* 39, 60–71.

GURLITT, J. AND RENKL, A. 2010. Prior knowledge activation: How different concept mapping tasks lead to substantial differences in cognitive processes, learning outcomes, and perceived self-efficacy. *Instructional Science 38,* 4, 417–433.

HAN, J. AND KAMBER, M. 2010. *Data Mining: Concepts and Techniques*, 2 ed. The Morgan Kaufmann series in data management systems. Elsevier/Morgan Kaufmann, Amsterdam.

HAY, D. B. AND KINCHIN, I. M. 2006. Using concept maps to reveal conceptual typologies. *Education + Training 48,* 2/3, 127–142.

HOFFMAN, R. R., SHADBOLT, N. R., BURTON, A. M., AND KLEIN, G. 1995. Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes 62,* 2, 129–158.

KARAGIORGI, Y. AND SYMEOU, L. 2005. Translating constructivism into instructional design: Potential and limitations. *Educational Technology & Society 8,* 1, 17–27.

KAUFMAN, L. AND ROUSSEEUW, P. J. 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley-Interscience paperback series. Wiley, Hoboken and N.J.

KEPPENS, J. AND HAY, D. B. 2008. Concept map assessment for teaching computer programming. *Computer Science Education 18,* 1, 31–42.

KINCHIN, I. M. 2000. Concept mapping in biology. *Journal of Biological Education 34,* 2, 61–68.

KINCHIN, I. M. AND CABOT, L. B. 2009. An introduction to concept mapping in dental education: the case of partial denture design. *European Journal of Dental Education 13,* 1, 20–27.

KINNEBREW, J. S., SEGEDY, J. R., AND BISWAS, G. 2014. Analyzing the temporal evolution of students' behaviors in open-ended learning environments. *Metacognition and Learning 9,* 2, 187–215.

KOPONEN, I. T. AND PEHKONEN, M. 2010. Entropy and energy in characterizing the organization of concept maps in learning science. *Entropy 12,* 7, 1653–1672.

KWON, S. Y. AND CIFUENTES, L. 2009. The comparative effect of individually-constructed vs. collaboratively-constructed computer-based concept maps. *Computers & Education 52,* 2, 365–375.

LARRAZA-MENDILUZE, E. AND GARAY-VITORIA, N. 2013. Use of concept maps to analyze students' understanding of the i/o subsystem. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research, Koli, Finnland, 14-17 November 2013*, M.-J. Laakso and Simon, Eds. Koli Calling '13. ACM, New York, 67–76.

LEAKE, D. B., MAGUITMAN, A., AND REICHHERZER, T. 2005. Understanding knowledge models: Modeling assessment of concept importance in concept maps. In *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society, Chicago, USA, 4-7 August 2004*, K. Forbus, D. Gentner, and T. Regier, Eds. Lawrence Erlbaum Associates, Mahwah and N.J, 785–800.

LEAKE, D. B., REICHHERZER, T., CAÑAS, A. J., CARVALHO, M., AND ESKRIDGE, T. 2004. "googling" from a concept map: towards automatic concept-map based query formation. In *Concept Maps: Theory, Methodology, Technology*, A. J. Cañas, J. D. Novak, and F. M. Gonzlez Garca, Eds. Vol. 1. 409–416.

MADHYASTHA, T. AND HUNT, E. 2009. Mining diagnostic assessment data for concept similarity. *Journal of Educational Data Mining 1,* 1, 72–91.

MCCLURE, J. R., SONAK, B., AND SUEN, H. K. 1999. Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching 36,* 4, 475–492.

NOVAK, J. D. 2010. *Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*, 2 ed. Routledge, London.

NOVAK, J. D. AND CAÑAS, A. J. 2008. The theory underlying concept maps and how to construct and use them: Technical report ihmc cmaptools 2006-01 rev 01-2008,.

NOVAK, J. D. AND CAÑAS, A. J. 2010. The universality and ubiquitousness of concept maps. In *Concept Maps: Making Learning Meaningful*, J. Sánchez, A. J. Cañas, and J. D. Novak, Eds. Vol. 1. Universidad de Chile, Chile, 1–13.

NOVAK, J. D. AND MUSONDA, D. 1991. A twelve-year longitudinal study of science concept learning. *American Educational Research Journal 28,* 1, 117–153.

OECD. 2012. *PISA 2009 Technical Report*. PISA. OECD Publishing, Paris.

OZDEMIR, A. 2005. Analyzing concept maps as an assessment (evaluation) tool in teaching mathematics. *Journal of Social Sciences 1,* 3, 141–149.

R CORE TEAM. 2013. R: A language and environment for statistical computing.

ROSAS, S. R. AND KANE, M. 2012. Quality and rigor of the concept mapping methodology: A pooled study analysis. *Evaluation and Program Planning 35,* 2, 236–245.

RUIZ-PRIMO, M. A. 2004. Examining concept maps as an assessment tool. In *Concept Maps: Theory, Methodology, Technology*, A. J. Cañas, J. D. Novak, and F. M. Gonzlez Garca, Eds. Vol. 1. 555–562.

RUIZ-PRIMO, M. A. AND SHAVELSON, R. J. 1996. Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching 33,* 6, 569–600.

SABITZER, B. 2011. Neurodidactics: Brain-based ideas for ict and computer science education. *The International Journal of Learning 18,* 2, 167–177.

SANDERS, K., BOUSTEDT, J., ECKERDAL, A., MCCARTNEY, R., MOSTRÖM, J. E., THOMAS, L., AND ZANDER, C. 2008. Student understanding of object-oriented programming as expressed in concept maps. *ACM Inroads 40,* 1, 332–336.

SCHVANEVELDT, R. W., DURSO, F. T., AND DEARHOLT, D. W. 1989. Network structures in proximity data. *The Psychology of Learning and Motivation 24*, 249–284.

SCHWARZ, G. 1978. Estimating the dimension of a model. *Annals of Statistics 6,* 2, 461–464.

SHAW, M. L. G. AND WOODWARD, J. B. 1990. Modeling expert knowledge. *Knowledge Acquisition 2,* 3, 179–206.

SOLOMON, K. O., MEDIN, D. L., AND LYNCH, E. 1999. Concepts do more than categorize. *Trends in cognitive sciences 3,* 3, 99–105.

SOUSA, D. A. 2009. *How the Brain Learns: A Multimedia Kit for Professional Development*, 3 ed. Corwin Press, Thousand Oaks and Calif.

SQUIRE, L. R. 1987. *Memory and Brain*. Oxford University Press, New York.

STIBOR, T. 2008. Discriminating self from non-self with finite mixtures of multivariate bernoulli distributions. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, Atlanta, USA, July 12-16, 2008*. GECCO '08. ACM, New York.

TARICANI, E. M. AND CLARIANA, R. B. 2006. A technique for automatically scoring open-ended concept maps. *Educational Technology Research and Development 54,* 1, 65–82.

TIBSHIRANI, R., WALTHER, G., AND HASTIE, T. 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63,* 2, 411–423.

TRUMPOWER, D. L. AND GOLDSMITH, T. E. 2004. Structural enhancement of learning. *Contemporary Educational Psychology 29,* 4, 426–446.

TRUMPOWER, D. L., SHARARA, H., AND GOLDSMITH, T. E. 2010. Specificity of structural assessment of knowledge. *The Journal of Technology, Learning, and Assessment 8,* 5.

VALERIO, A., LEAKE, D. B., AND CAÑAS, A. J. 2008. Automatic classification of concept maps based on a topological taxonomy and its application to studying features of human-built maps. In *Concept Mapping: Connecting Educators*, A. J. Cañas, P. Reiska, M. Åhlberg, and J. D. Novak, Eds. Vol. 1. Tallinn University, Estonia, 122–129.

VILLALON, J. J. AND CALVO, R. A. 2008. Concept map mining: A definition and a framework for its evaluation. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Sydney, Australia, 9-12 December 2008*. Vol. 3. IEEE, 357–360.

VON GLASERSFELD, E. 1989. Constructivism in education. In *The International Encyclopedia of Education*, T. Husén and T. N. Postlethwaite, Eds. Pergamon, Oxford, 162–163.

WITTROCK, M. C. 1992. Generative learning processes of the brain. *Educational Psychologist 27,* 4, 531–541.

WOLFE, J. H. 1970. Pattern clustering by multivariate mixture analysis. *Multivariate Behavioral Research 5*, 329–350.

YIN, Y., VANIDES, J., RUIZ-PRIMO, M. A., AYALA, C. C., AND SHAVELSON, R. J. 2005. Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching 42,* 2, 166–184.

YOO, J. S. AND CHO, M.-H. 2012. Mining concept maps to understand university students' learning. In *Proceedings of the 5th International Conference on Educational Data Mining, Chania, Greece, 19-21 June 2012*, K. Yacef, O. Zaïane, H. Hershkovitz, M. Yudelson, and J. Stamper, Eds. 184–187.