Concept systematization with concept maps in data modelling

MÁRTA CZENKY and JÁNOS KORMOS

Abstract. An important goal of concept learning is that students can allocate concepts in the hierarchical system of concepts. In the data modelling course, first, we supported concept systematization with worksheets in which the students had to fill in the blank hierarchical figures of classification of the concepts or blank Venn diagrams describing the relationships between concepts. The hierarchical systems, however, are somewhat restricted to the description of connections. The filling in Venn diagrams did not deliver the expected result, so our attention turned to concept maps. In this paper we introduce the concept maps we drew. Then we evaluate the results of concept mapping survey conducted among students. The survey was done in three courses. We compare the results of our survey with the result of an earlier concept systematising survey.

Key words and phrases: data modelling, concept learning, concept systematization, concept hierarchy, concept map.

ZDM Subject Classification: P20, Q50, Q60, R50, U60.

1. Introduction

The concept map is a graphical tool which shows the inner connections between concepts. Its important features are completeness and transparency. The use of concept maps facilitates systematization and organization of the information and is conducive to efficient learning. By using the concept maps the knowledge of the students gets more stable and organised.

The publication was supported in part by the TÁMOP-4.2.2C-11/1/KONV-2012-0001 project supported by the European Union, co-financed by the European Social Fund.

Copyright © 2014 by University of Debrecen
Since the establishment of concept maps (Novak) in the early 1970s they have been applied in a number of areas both in secondary and higher education. One can use the concept maps, among others, for summary and/or organization of knowledge, problem solving and systematization.

In the data modelling and database management courses the concept maps were used for describing data models, structure of tables, and connections between database management concepts.

In our research we designed the concept maps of data modelling related to entity-relationship and relational models, which we use primarily for concept systematization.

At the Faculty of Mechanical Engineering and Faculty of Agricultural and Environmental Sciences of Szent István University we teach database management for mechanical engineer, engineering manager and environmental engineer students on both BSc and MSc levels. We summarized the characteristics of the taught subjects in Table 1.

The curriculum of database management is the same in the case of all four subjects. In near equal (time) ratio we teach data modeling (E-R and relational model, normalization) and SQL-92 language [4], [17]. At the four classes per week courses students attend database management in the first half of the semester, while in the second half of the semester the KM3 students learn CAD and the KDB students learn about environmental databases.

Table 1. The characteristics of the courses

<table>
<thead>
<tr>
<th>Name of subject</th>
<th>Abbreviation</th>
<th>Major</th>
<th>Level/ Course</th>
<th>Class per week lecture + practise</th>
<th>Year</th>
<th>Head count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Informatics</td>
<td>MM</td>
<td>engineering manager</td>
<td>MSc, correspondence</td>
<td>0 + 2</td>
<td>2013</td>
<td>25</td>
</tr>
<tr>
<td>Computer Studies III</td>
<td>KM3</td>
<td>environmental engineer</td>
<td>BSc, full time</td>
<td>2 + 2</td>
<td>2013</td>
<td>37</td>
</tr>
<tr>
<td>Database Management</td>
<td>ABK</td>
<td>mechanical engineer</td>
<td>BSc, full time</td>
<td>0 + 2</td>
<td>2010</td>
<td>14</td>
</tr>
<tr>
<td>Environmental Databases</td>
<td>KDB</td>
<td>environmental engineer</td>
<td>MSc, full time</td>
<td>2 + 2</td>
<td>2013</td>
<td>9</td>
</tr>
</tbody>
</table>

Firstly we conducted a concept systematiser survey at the ABK course. The worksheet consisted of blank figures for hierarchical classification of the concepts, blank Venn diagrams for the description of the relationship of the concepts and
blank tables for the description of connections of concepts related to table concept of relational model. The students had to fill in the blank figures and tables.

14 students attended the course, but only ten of them participated in the survey. The random pattern involves these 10 students. Considering marks at the end of the semester by a homogeneity test the pattern and the course have same distribution.

The tasks of concept systematizing survey of ABK course were the same than the tasks of 2013 survey, but the students had to fill hierarchical figures and Venn-diagrams. We could not repeat this survey because the teaching of Database management subject was suspended for mechanical engineering student in recent years.

The hierarchical systems are somewhat restricted to the description of the connections. Filling in Venn diagrams and tables did not deliver the expected results. The average result was 63.9%, which is a very weak satisfactory result. Therefore our attention turned to concept maps and we drew the concept maps of data modelling. The concept maps include more information than hierarchies and Venn diagrams, for example through labels written on connective lines. These items of information support recognition of connections between concepts and concept systematization.

At the MM, the KDB and the KM3 courses we executed a survey in 2013, the students had to fill in the worksheets at the exam. All students of the courses participated in this survey. We evaluate the results of the survey later below.

2. Learning Concepts

In this section we review those criteria of concept learning which are relevant to our research.

“The concept is a mental form developed from the main characteristics of the things in the consciousness” [8], [11]. In the course of the constitution of a concept we unite the essential characteristics into a standardized form, and we disregard the irrelevant features. The content of the concept is the totality of the substantive characteristics; the extent of the concept is the totality of all things which dispose the substantive characteristics creating the concept.

There are several ways of introducing concepts:

• the inductive way, which is getting through abstraction from the comparison of concrete examples to the general,
• the deductive way, which is getting from general to concrete,
• the constructive way, producing a representative of the concept under given conditions, then generalizing the procedure.

We may consider the concept is acquired if the student can give the definition of the concept, can identify the concept, can list examples (concept realization), knows the characteristics of the concept, can place the concept into the hierarchical system of the concepts moreover is able to use the concept for problem solving [1], [6].

The ability to systematise is instrumental in the exploration of the connection existing between concepts and configuration of their hierarchical system. According to [13] definition the systematising capability makes it possible to create new knowledge by the recognition and ordering of the things and their relationships or rather the extant information and their relationships.

In the course of hierarchical classification we work deal with the hierarchy reflecting the relationship of concepts. The concepts may be in subordinate, superior or coordinate contact with each other. The subordinate concept is at the same time a superior concept, too. We come to the superior concept with the omission of a substantive characteristic of the subordinate concept. Between the coordinate concepts there are no subordinate-superior relations, but they have a common superior concept [1].

Inadequate knowledge of the concept systems obstructs recognition of things and phenomena, permanent knowledge fixing and the use of concepts in problem solving.

In general, among concepts we can distinguish objective, relational, and operational concepts. This classification of concepts appears in the classification of data modelling concepts, too.

3. Categories and teaching data modelling

The learning of concepts and the exploration of the relationship between the concepts are necessary terms of database design and database management. [16] analyzed the database planning concept of seventeen database course books, with regard to the importance of concept learning in database planning. He emphasizes the importance of correct concept identification and the knowledge of the characteristics of the concepts, which are necessary for the categorization of fundamental concepts. [3] also emphasize the importance of these two steps
of concept learning in connection with teaching the E-R model, although their article primarily deals with the use of UML technology in data modelling.

In the field of data modelling, besides the fundamental concepts, the entity relationship model, the relational model, the formulation of entity relationship diagrams, dependence diagrams and normalization as database planning method are also taught. Based on these we can classify the concepts of data modelling as follows:

- The basic concepts of data modelling: entity and set of entities, relationship and occurrence of relationship, attribute and value of attribute. The data models give mappings of these abstract and concrete concepts into certain structures, so these concepts also appear in data models, but they are not listed separately there.
- Entity-relationship model (E-R): hierarchy of set of entities, ISA (from “is a” combination) relationship, inheritance of attributes, integrity constraints.
- Relational model: relation (relational table), keys, special data values (Null, default), integrity constraints and their enforcements, index, view, synonym, relational operations.
- Dependence: functional dependence, multivalued dependence, and their types.
- Dependence diagrams: the depiction of relational dependences in a table on a diagram.
- Normalization: closing process, normal forms [4], [17].

Among the concepts we find objective concepts (for example entity, set of entities, relation, primary key, etc.), relational concepts (for example relationship, occurrence of relationship, etc.) and operational concepts (for example relational operations, cascade update and deletion, etc.) simultaneously.

The tasks of a printed exercise book and questions of tests in Moodle course management system help concept learning. The measurement of the acquirement of the concepts is done with electronic exam tests, while we measure the students’ capability of concept systematization in writing.

The tests and the tasks of exercise book check the knowledge of the definitions, the identification of the concepts and the knowledge of the characteristics of the concepts. Although in tests of the Moodle system the evaluation of the answer is automatic this system enables concept realization only with direct evaluation of the answer by the teacher. Therefore we always request the students to list the examples during the class. Concept realisation tasks are the ones in which you
have to recognize examples given for a certain concept. There are some smaller modelling tasks, too, which make the application of the knowledge of the concepts possible in problem solving [5].

In data modelling we have introduced two new steps of concept learning: the mapping of the concepts of different model and the mapping of the concepts of different depiction. For the first one the mapping of the concepts of E-R and relational model can be an example, while for the second the mapping of the concept of relational model is described by schema and dependence diagram.

4. Concept map

The concept map is a diagram which shows the inner relationships among concepts. It is a graphical tool of organising and visualising knowledge. The concept map was developed by Novak in the 1970s. The concept map helps to understand concepts, phenomena and processes [15].

By [7] the use of concepts map supports and promotes the exploration of concepts, thinking in processes, problem solving and information recall.

The concept map is a type of graphical organizer tool, which denotes the concept with labelled nodes and the relationship among the concepts with lines. The lines of the concept map may be labelled or unlabelled, directional or non-directional [14].

[15] describes the concept map as follows:

• the concept map includes concepts and relationships among concepts,
• in the concept map the concepts usually appear in some kind of circles or boxes,
• the label assigned to the concept is mostly a word,
• a connecting line indicates the relationship among the concepts,
• words on connecting lines determine the relationships,
• the concept map is a hierarchical structure, too, the general concept appears on the top of the concept map, while the less general concept at its bottom,
• cross-links are possible on the concept map, which denote relationship among different areas of the concept map,
• the concept map can include events and objects, too, which help to clarify the meaning of the concept, we do not enclose them in circles or boxes.
The concept maps can be used widely for summation in education, organization of knowledge and problem solving. There are three main parts of its use:

- supporting learning including assessment,
- directional teaching, designing educational content and guideline tool,
- organising and introducing information [2].

There are instances for using the concept maps in data modelling, database design and database management, too. [9] describes data models with concept maps then transcribes them into entity-relationship diagrams. [10] introduces how to describe structure and relationship of tables of relational model with concepts maps. [12] uses concept maps for teaching concepts of distributed databases.

5. Concept maps of data modelling

For the sake of a more exact and more complete overview of the connection between data modelling concepts we created the data modelling concept maps introduced in this section. These concept maps describe the connections of the concepts of entity-relationship and relational model. We intend to use these concepts maps for the systematization and deepening of the students' knowledge.

![Figure 1. Connections between basic data modeling concepts](image-url)
The first four figures show the connections of the concepts of the E-R model. Figure 1 shows the connections between basic data modeling concepts. Among the concepts appearing in the figure the set of entities, the relationship and the attribute are general concepts, while the entity, the occurrence of relationship and the value of attribute are concrete concepts. In database management literatures the usual name of general concepts is entity type, relationship type and attribute type, while the concrete concepts are called occurrences, too.

The figure carries structural information (consists of), connecting information (belong to), descriptive information (characterizes) and containing information (value of). You cannot visualize these connections in a hierarchical system because this classification is not hierarchical.

Figure 2 describes the classification of the relationships. The classification can be done by two points of view:

- by type – the cardinality of the relationship,
- by the number of the participating set of entities - degree of relationship.

![Classification of relationships](image)

Figure 2. Classification of relationships

In a hierarchical system this means drawing two hierarchies, while the concept map simultaneously includes the classification by both points of view. This corresponds to practice better because we usually classify the relationships by
both points of view at the same time. The figure is also suitable for the indi-
cation of such information that we classify by type only between entities of one
or two sets of entities. In the classification by the number of participating set of
entities appear the general concepts as set of entities, while in classification by
the type appear the relationships between the concrete concepts as entities.

Figure 3. Classification of attributes

Figure 3 shows the classification of the attributes which we do simultaneously
by several points of view. The classification can happen
• by construction (simple, composite)
• by the number of data values (single-valued, multi-valued),
• by identifier capability (identifier, non-identifier),
• by the source of data values (derived, non-derived).

In a hierarchical system the four kinds of classification means the drawing of
four hierarchies, while in concept map you can visualize each item of information
simultaneously. The figure does not include the non-identifier and non-derived
attributes because in classification we use only identifier and derived concepts.
In classification by value you can visualize Null value but we teach this concept
with relational model, therefore we do not feature it in the figure.
Figure 4. Classification of sets of entities

The Figure 4 shows the classification of the sets of entities. This concept map is almost a hierarchy; the difference is that the figure includes the points of view of the classification.

Figure 5. Classification of keys

The 5th, 6th and 7th figures relate to relational model.

On Figure 5 you can see the classification of the keys. The left and the right arm are practically a hierarchy with indication of the points of view of the classification. In the classification of the keys we generally disregard the indication
of foreign key and grant only the keys playing a part in the identification. Due to its importance we marked the foreign key in the concept map. We generally use the distinction by construction (elementary, composite) only for primary keys. However, as the foreign key is always connected with some primary key we applied the distinction also for foreign key.

![Diagram of normal forms](image)

**Figure 6. Connections between normal forms**

Figure 6 describes the connection of normal forms. It is usually depicted by embedded sets indicating that only fewer tables satisfy the higher normal forms. Our solution includes this information, too, but representation with concept map is better because in this way the figure can include those conditions which ensure the fruition of normal forms and the arrows signify the procession of normalisation.

Figure 7 shows the connections between concepts connected to the table concept of the relational model. This figure carries information
- about the structure of the table,
- about the things described by some elements of the table,
- about the content of the elements,
- about the identification of the elements of the table,
• about integrity constraints relating to primary and foreign keys,
• about constraints relating to the elements of the table.

6. The concept systematising survey and its assessment

In the first semester of the 2013/2014 school year we conducted the survey with concept maps among the students of the MM, the KDB and the KM3 courses. The students had to fill in the questionnaire at the exam. All the students of the courses participated in the survey, so the result of the survey can be considered as representative.

The worksheet included nine tasks, see Table 2. The name of the correct concepts had to be written into blank rectangles of the figures of the worksheet. The students did not see the figures before the exam. By facilitation we listed

Figure 7. The concept map of the table of the relational model

From among the concept maps we created this is the most complicated expectedly filling in this one will cause the most problems for the student.
the concepts on the worksheet. We wrote the names of the concepts into some rectangles in advance which are signed in boldface type in Figures 1–7.

Table 2. The tasks of the concept systematizing survey in 2013

<table>
<thead>
<tr>
<th>Number of task</th>
<th>Number of figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>determination of connection between basic data modelling concepts</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>classification of the relationships</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>classification of the attributes</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>classification of the sets of entities</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>classification of the keys</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>determination of connection between concepts related to table concept of relational model</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>classification of the dependencies, hierarchical figure</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>determination of connection between normal forms</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>classification of the operations, hierarchical figure</td>
</tr>
</tbody>
</table>

As teachers we considered that the most difficult task will be to fill in Figure 7. Therefore we summarized the connections of the concepts related to the table concept of the relational model in a table, see Table 3.

Table 3. The connections of the concepts related to the table concept of the relational model

<table>
<thead>
<tr>
<th>Part</th>
<th>Content</th>
<th>Emphasised role</th>
<th>Constraint for content</th>
<th>Identification – Integrity constraint</th>
<th>Description of relationship – Integrity constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table</td>
<td>Set of entities and/or relationship</td>
<td></td>
<td></td>
<td>Table name – unique</td>
<td></td>
</tr>
<tr>
<td>1st row</td>
<td>Attributes (columns)</td>
<td>Primary key</td>
<td>Different</td>
<td>Attribute name (column name) – unique</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>Entity and/or occurrence of relationship</td>
<td>Different</td>
<td></td>
<td>Values of primary key – value based identification</td>
<td>Value of foreign key – value based description of relationship</td>
</tr>
<tr>
<td>Fields</td>
<td>Value of attribute</td>
<td>Atomic value</td>
<td>Unique value, not Null</td>
<td>Reference condition</td>
<td></td>
</tr>
</tbody>
</table>

While correcting the worksheets we collected the typical errors by tasks, which were the following:
With giving the connections of the basic data modelling concepts (Figure 1) the most frequent error was the confusion of the general and concrete concepts, for example the entity consists of sets of entities. It is interesting that everybody gives the connection of attribute and value of attribute correctly. We think that the names of the concepts helped with it.

With the exception of a few all the students filled in the concept map of the classification of the relationships properly (Figure 2). In the incorrect solutions the students confused the classifications by type and by the number of participation.

The majority of the students solved properly the filling of the concepts map describing the classification of the attributes (Figure 3). Those who did not succeed in filling it in made the following mistakes:

- the attribute can be identifier and derived by construction,
- the simple attribute identifies the entity,
- the composite attribute derives from other attributes.

The student who made an error in the classification of the sets of entities confused the existence dependent set of entities and weak sets of entities (Figure 4).

In the classification of the keys (Figure 5) the most frequent error was the confusion of the primary key and super key. Besides this error confusion of elementary key and foreign key or even super key and alternative key also occurred.

In giving the connections of the concepts related to table concept of the relational model (Figure 7) it was a frequent mistake that the column describes the name of attribute and the 1st row consists of attributes. Sometimes students mixed up the concepts described by table or by row. In addition it also occurred that the tuple describes foreign key and the name of attribute identifies the table.

In the classification of the dependencies the most frequent error was that students confused the functional and multivalued dependencies. Besides this it occurred that the multivalued dependency appeared as a type of functional dependency. It was an error, too, that students gave the full and partial dependencies as a type of multivalued dependency.

In the correspondence of normal forms (Figure 6) some students confused the BCNF and 4th normal form.
• The errors occurring in the classification of the operations:
  
  – the lack of two tables operation,
  
  – the students are not aware of the three fundamental two tables operation type (Cartesian product of rows, join, set operation) and their types,
  
  – they do not recognize the connection of the join and natural join.

Considering the result of the survey in the case of all three courses the classification of the relationships and attributes and the giving of connections between normal forms went in the most flawless way, the students achieved circa 90% result in these tasks. The students achieved the worst result in filling Figure 7 and the classification of the dependencies and operations tasks. In other tasks the results were satisfactory.

We experienced in four tasks that the students left the figures untouched. One student did not fill in the figure describing the connection of the basic data modelling concepts. In KM3 course 9 students, in KDB course 2 students and in MM course 1 student did not fill in Figure 7. 12 students (KM3), 2 students (KDB) and 3 students (MM) did not fill in the figure of the table operations. We consider these values to be high. We think that the reason for the blank Figure 7 is the difficulty of the task, while the reason for the blank figure of table operations is that students did not learn the operations.

The ratio of the students who solved the task correctly verifies the statements written previously. In KM3 course not more than 3 students, in KDB course 1 student and in MM course 9 students were able to fill in Figure 7 without any mistake. Obviously this task was the most difficult for the students.

In KM3 course 6 students were able to give the table operation correctly while in the other two courses nobody was able to do it. The third most complicated task was the giving of the connections of the basic data modeling concepts. 9 (KM3), 5 (KDB) and 9 (MM) students solved it properly. We regard this value as low because the acquirement of the basic modelling concepts is the root of data modelling, you can’t model without the knowledge of these concepts.

We examined whether the results show normal distribution. The results of the KM3 and KDB courses do not have normal distribution but the result of MM courses does. The result of ABK courses has normal distribution, too.

We made homogeneity tests between the results of ABK course and the courses participated in 2013 survey to decide if the distributions of the results are the same. Because not all of results have normal distribution we could not
execute F-test and T-test in every case. Instead of it we examined the homogeneity of standard deviations with Levene-test and homogeneity of means with Mann–Whitney test.

Table 4. The results of homogeneity tests

<table>
<thead>
<tr>
<th></th>
<th>ABK-KM3</th>
<th>ABK-KDB</th>
<th>ABK-MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>standard deviation</td>
<td>Levene-test</td>
<td>Levene-test</td>
<td>F-test</td>
</tr>
<tr>
<td></td>
<td>not same</td>
<td>not same</td>
<td>same</td>
</tr>
<tr>
<td>mean</td>
<td>Mann-Whitney test</td>
<td>Mann-Whitney test</td>
<td>two samples T-test</td>
</tr>
<tr>
<td></td>
<td>not same</td>
<td>same</td>
<td>not same</td>
</tr>
</tbody>
</table>

We summarized the executed tests and their results in Table 4. The results show that the distributions are not the same in any case. Comparing the result of 2013 surveys with the result of 2010 survey we see that the average scores achieved in 2013 survey were higher (KM3-6.46, KDB-6.78, MM-6.67) than in the survey conducted in 2010 (ABK-5.75). As the standard deviation and mean of KM3 course are not the same as the similar data of ABK course and the mean of KM3 is better, we can say that the result of KM3 is significantly better than the result of ABK course with 0.05 significance level.

Because the four patterns are independent, in all three courses of 2013 survey we did an independence test too to decide whether the result of the concepts mapping survey is significantly better than the result of concept systematiser survey done in 2010. According to this the result of KM3 course is significantly better with 94% probability than the result of the survey conducted in ABK course in 2010. We cannot claim the same in KDB and MM courses.

7. Summary

All in all we can state that the construction of the concept maps and using them for concept systematisation in education proved to be a good practice.

The use of concept maps has several advantages compared to the hierarchical classification and classification with Venn diagram:

- in hierarchical system you cannot describe some connections because they do not describe the subordinate, superior or coordinate contact of concepts, for example between basic concepts of data modelling or between concepts related to table concept of relational model,
you can classify some concepts by several aspects, but you can only describe them with more individual hierarchy, for example attributes, while with concept maps you can describe the classification by several aspects,

- the advantage of concept map is that the view-points of the classification appear on them,

- the concepts maps can include information which cannot appear in hierarchical classification, for example the restrictions done for the concepts. (See Figure 7.)

With using concept maps the average scores achieved in this survey were higher than in the survey conducted in 2010. We could also verify at one course that the results were significantly better in the concept mapping survey.

Assessing the results we can say that those students who achieved better result than 76% learned concepts adequately and they recognised the connection between concepts. We consider acceptable the knowledge and concept systematiser capability of the students who achieved satisfactory result, while we need to help the concept learning of students who achieved worse result than satisfactory with different tools.

References


MÁRTA CZENKY
DEPARTMENT OF INFORMATICS
FACULTY OF MECHANICAL ENGINEERING
SZENT ISTVÁN UNIVERSITY
1084 BUDAPEST, DÉRI MIKSA UTCA 15. 3/5
HUNGARY
E-mail: marta.czenky@t-online.hu

JÁNOS KORMOS
DEPARTMENT OF APPLIED MATHEMATICS AND PROBABILITY THEORY
FACULTY OF INFORMATICS
UNIVERSITY OF DEBRECEN
4010 DEBRECEN, P.O. BOX 12
HUNGARY
E-mail: kormos.janos@inf.unideb.hu

(Received January, 2014)