Knowledge Evaluation with Rough Sets

Sylvia Encheva[1,2(✉)] and Torleiv Ese[1]

[1] Stord/Haugesund University College,
Bjørnsonsg. 45, 5528 Haugesund, Norway
{sbe,te}@hsh.no
[2] Polytec, Sørhauggata 128, 5527 Haugesund, Norway

Abstract. Students often experience significant difficulties while studying mathematical subjects. In this work we focus on a course in calculus given to third year bachelor engineering students. The course is optional with respect to completing a bachelor degree and compulsory for taking a master degree in engineering. Our intention is to find out whether early identification of students in danger to fail the subject is possible and if affirmative which factors can be used to support the identification process. Methods from rough set theory are applied for selection of important attributes and factors influencing learning.

Keywords: Rough Sets · Logic · Knowledge Evaluation

1 Introduction

A number of students have serious difficulties studying mathematics at all levels. Researches from different fields have been working on this problem for over a century. In [6] we find a nonexhaustive list including cognitive factors, metacognitive factors, habits of learning and previous experiences related to studying mathematics. In [8] it is pointed out that if a student is unable to understand a difficult mathematics class because it is at a level above student's ability to respond to the instruction, the student may not progress to the affective level of valuing the instruction.

In this work we consider problems which third year bachelor engineering students have while completing a course in mathematics that is optional with respect to obtaining a bachelor degree and a prerequisite for taking a master degree in engineering. Our intention is to find out whether early identification of students in danger to fail their final exam is possible and which factors can be used in that identification process.

The volume of the course corresponds to one third of a study load per semester. Students are offered fourteen weeks face to face classroom lectures and tutorials. The former takes four hours a week and the latter is two hours a week. Students have to deliver one course work, take a two hours midterm exam, and a five hours final exam. Both the midterm exam and the final exam are written and taken in a controlled location. The course work is to be completed at home. This means that they can collaborate even though they are supposed to work individually. The course work is marked as 'pass' or 'fail'.

Methods from rough sets theory [10, 11] are employed in the course of our investigations.

© Springer International Publishing Switzerland 2015
D.-S. Huang and K. Han (Eds.): ICIC 2015, Part III, LNAI 9227, pp. 179–186, 2015.
DOI: 10.1007/978-3-319-22053-6_20

The rest of the paper goes as follows. Theoretical supporting the study is presented in Sect. 2, our approach can be found in Sect. 3 while the conclusion is placed in Sect. 4.

2 Preliminaries

Rough Sets were originally introduced in [10]. The presented approach provides exact mathematical formulation of the concept of approximative (rough) equality of sets in a given approximation space.

An *approximation space* is a pair A = (U, R), where U is a set called universe, and $R \subset U \times U$ is an indiscernibility relation [11].

Equivalence classes of R are called *elementary sets* (atoms) in A. The equivalence class of R determined by an element $x \in U$ is denoted by $R(x)$. Equivalence classes of R are called *granules* generated by R.

The following definitions are often used while describing a rough set $X, X \subset U$:

the R-*upper approximation* of X

$$R^*(x) := x \in UR(x) : R(x) \cap X \neq \emptyset\},$$

the R-*lower approximation* of X

$$R_*(x) := x \in UR(x) : R(x) \subseteq X\},$$

the R-*boundary region* of X

$$RN_R(X) := R^*(X) - R_*(X).$$

Elements in the index set $A = a_1, a_2, \ldots, a_m$ are the importance degree of attribute set where each index in the system is determined by:

$$S_A(a_i) = \frac{|POS_A(A)| - |POS_{A-a_i}(A)|}{|U|}$$

where $i = 1, 2, 3, \ldots, m$ and the weight of index a_i is given by

$$w_i = \frac{S_A(a_i)}{\sum_{i=1}^{m}(a_i).}$$

The assessment model is defined by

$$P_j = \sum_{i=1}^{m} f_i$$

where P_j is the comprehensive assessment value of assessed jth object, f_i is the assessment value of i th index a_i according to the comprehensive assessment value,

[12]. Rough set theory software can be downloaded from [7]. Attributes degree of importance is discussed in [15].

Let P be a non-empty ordered set. If $sup\{x, y\}$ and $inf\{x, y\}$ exist for all $x, y \in P$, then P is called a *lattice* [5]. In a lattice illustrating partial ordering of knowledge values, the logical conjunction is identified with the meet operation and the logical disjunction with the join operation.

A *context* is a triple (G, M, I) where G and M are sets and $I \subset G \times M$. The elements of G and M are called *objects* and *attributes* respectively, [16].

For $A \subseteq G$ and $B \subseteq M$, define

$$A^{'} = \{m \in M | (\forall g \in A) gIm\},$$

$$B^{'} = \{g \in G | (\forall m \in B) gIm\}.$$

where $A^{'}$ is the set of attributes common to all the objects in A and $B^{'}$ is the set of objects possessing the attributes in B.

A *concept* of the context (G, M, I) is defined to be a pair (A, B) where

$$A \subseteq G,\ B \subseteq M,\ A^{'} = B \text{ and } B^{'} = A.$$

The *extent* of the concept (A, B) is A while its *intent* is B. A subset A of G is the extent of some concept if and only if $A^{''} = A$ in which case the unique concept of the which A is an extent is $(A, A^{'})$. The corresponding statement applies to those subsets $B \in M$ which is the intent of some concepts. The set of all concepts of the context (G, M, I) is denoted by $B(G, M, I)$. $\langle B(G, M, I); \leq \rangle$ is a complete lattice and it is known as the *concept lattice* of the context (G, M, I).

Students' prior mathematical knowledge are of utmost importance when it comes to building of higher order mathematical understanding. Students mathematical competencies are discussed in [14]. Some problems related to building mathematical concepts are listed in [3]. Various problems concerning development of higher level thinking are presented in [1, 2, 9, 14]. The authors state that students' higher level thinking occurs when students are exposed to active learning. The latter can take place when "educators must give up the belief that students will be unable to learn the subject at hand unless the teacher "covers it"", [9]. In [4] higher-order thinking is divided in three categories - higher-order thinking in terms of transfer, in terms of critical thinking, and in terms of problem solving.

3 Data Evaluation

Data used in this study is taken from students results over a period of three years where 115 undergraduate engineering students have been enrolled in a mathematical course. The amount of students taking the course is usually about 40 and varies slightly from year to year.

The course is based on the following chapters in [13]:

12. Vectors and the Geometry of Space,
13. Vector-Valued Functions and Motion in Space,
14. Partial Derivatives,
15. Multiple Integrals,
16. Integrals and Vector Fields.

To illustrate the approach we present data for about halve of the students being subjects of this study, see Tables 1 and 2. Under 'Gender' notations are 'f' for female and 'm' for male. Note that no other descriptions of gender are used in students register.

Table 1. Students results, part 1

	Gender	M 1	M 2	Test	G 1	G 2	G 3
S1	f	a	a	h	h	h	h
S2	f	a	a	a	l	h	h
S3	f	a	a	h	h	a	h
S4	m	h	h	h	h	h	h
S5	m	h	h	h	h	a	h
S6	m	l	l	a	l	l	l
S7	m	l	a	a	a	a	l
S8	m	a	a	a	a	a	l
S9	m	a	l	a	a	l	a
S10	f	a	l	h	a	a	h
S11	m	a	l	h	h	a	h
S12	f	a	a	a	l	l	l
S13	f	a	l	a	l	l	l
S14	m	l	a	a	l	l	l
S15	m	a	h	a	a	l	a
S16	f	l	l	h	h	l	a
S17	m	h	a	h	l	a	a
S18	f	h	h	h	a	a	l
S19	f	a	a	a	h	a	a
S20	m	l	l	a	l	l	l
S21	m	a	a	a	l	h	a
S22	f	a	a	h	a	h	a
S23	f	a	a	h	h	a	l
S24	m	l	l	h	h	a	a
S25	m	l	l	a	a	l	l
S26	f	a	a	h	a	a	h
S27	m	a	l	h	h	h	a
S28	m	a	a	a	h	a	a
S29	f	a	l	a	l	l	l

Table 2. Students results, part 2

	Gender	M 1	M 2	Test	G 1	G 2	G 3
S30	f	l	a	a	l	l	l
S31	f	h	h	h	h	h	a
S32	m	h	a	h	h	h	a
S33	f	l	a	h	h	h	a
S34	f	a	l	h	a	h	h
S35	m	a	a	h	h	h	h
S36	f	h	h	h	h	a	h
S37	f	a	a	h	h	a	h
S38	m	a	h	h	h	a	h
S39	m	l	l	a	l	l	l
S40	f	h	h	h	h	h	h
S41	m	h	h	h	h	h	h
S42	m	h	h	h	h	h	a
S43	m	h	h	h	h	h	h
S44	f	h	h	h	h	h	a
S45	m	h	h	h	h	h	h
S46	f	h	h	h	h	h	h
S47	f	h	h	h	h	h	h
S48	m	h	a	h	a	a	h
S49	f	h	a	a	h	a	h
S50	m	h	h	h	h	a	h
S51	m	h	a	h	a	h	h
S52	f	a	a	h	h	a	h
S53	m	a	a	h	a	h	h
S54	f	l	l	h	l	l	l
S55	m	h	h	h	h	h	h

Final grades in two previously taken courses in mathematics at the same university are placed under columns ‘M1’ and ‘M2’, where:

– ‘M1’ stands for Mathematics 1 and
– ‘M2’ stands for Mathematics 2.
– Group 1 denoted ‘G1’ refers to application of triple integrals, moments and centers of mass, cylindrical and spherical coordinates;
– Group 2 denoted ‘G2’ refers to line integrals, surface integrals, Stokes’ theorem and Gauss theorem, and
– Group 3 denoted ‘sssG3’ refers to studying correlations between curves, surfaces and given functions.

Originally the data comes in both text and numerical form. The data in Tables 1 and 2 is converted to text form with scaling grades:

$\{A,B\}$ - high (h),
$\{C,D\}$ - average (a),

{E,F} - low (l);
tests and written evaluations
[0 %, 40 %] - low (l),
[41 %, 70 %] - average (a), and
[71 %, 100 %] - high (h).

Suppose students in that course belong to a rough set X. When the data set described as in Tables 1 and 2 is used we obtain the following distribution

$$R_*(X) = \{St1, \ldots, S5, S7, S10, S11, S18, S19, S22, S23, S24, S26, S27, S28, S31, \ldots, S38, S40, \ldots, S53, S55\},$$

$$RN_R(X) = \{S8, S9, S12, S15, S16, S17, S21, S25, S39\},$$

$$X - R_*(X) = \{St6, S13, S14, S20, S29, S30, S54\}$$

where $R_*(X)$ is the set of students who have obtained sufficient knowledge and skills, $RN_R(X)$ is the set of students who have obtained somewhat insufficient knowledge and skills, and the set of students who definitely have not obtained sufficient knowledge and skills is $X - R_*(X)$.

Below we present findings from working with data where all 115 students are included. Students who have not obtained sufficient knowledge and skills or have obtained somewhat insufficient knowledge and skills have lower grades from previously taken mathematical courses and their scores from the first test is in the interval [50 %, 60 %] nearly without exceptions. In the future such comparisons can be performed after the results from the first test are available and students who are in danger to fail the subject will be notified. This early warning will encourage them to spent more efforts on studying this subject for the rest of the semester.

The two groups of students,

$$RN_R(X) \text{ and } X - R_*(X),$$

have particular problems while working with line integrals, surface integrals, Stokes' theorem and Gauss theorem, as well as establishing correlations between curves, surfaces and given functions. An indept study of weaker students' performance indicates not only a gap in their mathematical knowledge but even more importantly, lack of higher order skills.

In general it seems that the majority of students have some problems with the above listed topics. However, students who have obtained sufficient knowledge and skills have scores on problem solving in the interval [60 %, 90 %] at average, while the group of students with insufficient or somewhat insufficient knowledge have scores in the interval [0 %, 30 %]. It turns out that very few students with pour grades from previously taken mathematical courses have managed to pass this course. Therefore, students, who for some reasons make slower progress than what previous experience indicates, should be notified as soon as possible and some further actions will be suggested to them.

In future work we will consider the following question: what can be done in the first two courses that will help students from potentially belonging to set $RN_R(X)$ to complete the course belonging to set $R_*(X)$, Fig. 1.

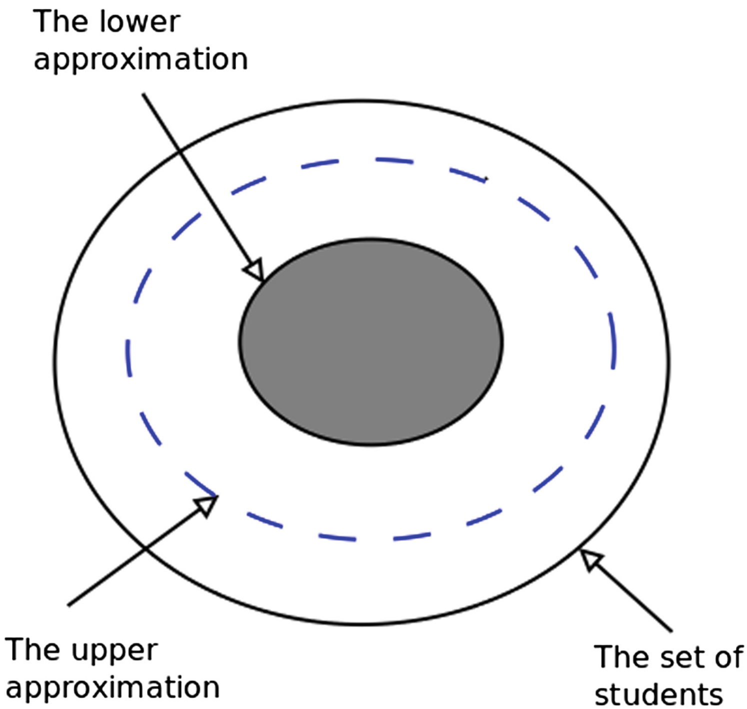

Fig. 1. Lower and upper approximations

A concept lattice can be depicted from data as presented in Tables 1 and 2 where fifteen students belonging to the three groups ‘high’, ‘average’, and ‘low’ are involved. The idea is to illustrate how formal concept analysis can be used in similar studies. Note that a lattice representing formal concepts for 115 students would need much larger space than the one we can use here. Concepts are presented in gray boxes where the lower row lists concepts objects and the upper row lists concepts attributes. When two concepts are connected with a straight line one can see that the lower concept contains less objects and more attributes while the upper concepts has less objects and more attributes.

Concepts in the lattice show which groups of students have the same results related to particular topics and where they differ in their performance. We can see what ‘high and average’ groups learn equally good and where exactly are the differences between ‘high’ and ‘average’ groups. This can support provision of tailored advise to students in group ‘average’ that will help them to join the ‘high’ group.

Similarly, students from the ‘low’ group can receive personal advise on how to join the ‘average’ group and possibly the ‘high’ group. Students with the same results related to particular topics can be found in concepts placed on the third and fourth rows. Their differences can be seen by following superconcepts placed in the fifth and sixth rows.

4 Conclusion

In this work we study correlations between students' knowledge obtained in preliminary courses in mathematics on undergraduate level and their results from a mathematical course on a lower graduate level. Their test results combined with previous history gives an indication on whether they are going to experience difficulties on some particular parts of the curriculum. Such tendencies can also be used to provide support to new students taking the same course.

References

1. Anderson, L.W., Krathwohl, D.R.: A taxonomy for learning, teaching, and assessing: a revision of bloom's taxonomy of educational objectives. Longman, New York (2001)
2. Ball, A.L., Garton, B.L.: Modeling higher order thinking: the alignment between objectives, classroom discourse, and assessments. J. Agric. Educ. **46**(2), 1–9 (2005)
3. Berger, M.: Making mathematical meaning: from preconcepts to pseudoconcepts to concepts. Pythagoras **63**, 14–21 (2006)
4. Brookhart, S.M.: How to Assess Higher-Order Thinking Skills In Your Classroom. Association for Supervision and Curriculum Development, New York (2010)
5. Davey, B.A., Priestley, H.A.: Introduction to lattices and order. Cambridge University Press, Cambridge (2005)
6. Gurganus, S.P.: Characteristics of Students' Mathematics Learning Problems. http://www.education.com/reference/article/students-math-learning-problems/
7. http://www.lcb.uu.se/tools/rosetta/
8. LeMire, S.D., Melby, M.L., Haskins, A.M., Williams, T.: The devalued student: misalignment of current mathematics knowledge and level of instruction. Math. Educ. **22**, 63–83 (2012)
9. Limbach, B., Waugh, W.: Developing higher level thinking. J. Instr. Pedagogies **3**, 1–9 (2010)
10. Pawlak, Z.: Rough sets. Int. J. Comput. Inform. Sci. **11**, 341–356 (1982)
11. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishing, Dordrecht (1991)
12. Ramasubramanian, P., Banu Priya, S., Dhanalakshmi, T.: Study on assessment method for technical institutions using rough sets. Int. J. Res Rev. Inf. Sci. (IJRRIS) **1**(4), 99–103 (2011)
13. Thomas Jr., G.B., Weir, M.D., Hass, J.R.: Thomas' Calculus, 13th edn. Pearson, New York (2014)
14. Turner R.: Exploring mathematical competencies, Research Developments **24**(5), 1–6 (2011). http://research.acer.edu.au/resdev/vol24/iss24/5
15. Yuan, L., Xu, F.: Research on the multiple combination weight based on rough set and clustering analysis-the knowledge transfer risk in it outsourcing taken as an example. Procedia Comput. Sci. **17**, 274–281 (2011)
16. Wille, R.: Concept lattices and conceptual knowledge systems. Comput. Math. Appl. **23**(6–9), 493–515 (1992)