KNOWLEDGE REPRESENTATION WITH AND-OR-GRAPHS: COMPARING THE APPROACH OF DOIGNON & FALMAGNE WITH THE ABSYNT-DIAGNOSTICS

Klaus Kohnert, Gabriele Janke
Project ABSYNT*, Computer Science Department, University of Oldenburg
P. O. Box 2503, D-2900 Oldenburg, West Germany

0. Abstract

Doignon & Falmagne as well as the ABSYNT project work in the field of Intelligent Computer-Aided Instruction. Both groups use AND-OR-graphs for knowledge representation. Doignon & Falmagne represent observable empirical data and use it for the assessment of the student's current knowledge state. The ABSYNT project represents the results of a task analysis of the given domain and uses it for the generation of adaptive helps. This report shows the different kinds of knowledge to be represented and the different applications. The comparison gives an impression of how wide the area of useful applications of AND-OR-graphs can be.

1. The Related Work of Doignon & Falmagne

1.1. Overview

The work of Doignon & Falmagne, as far as reported in [Doignon & Falmagne 1986], is aimed at providing a formalism which allows to assess a student's knowledge state with as few questions as possible. From the domain to be learned Doignon and Falmagne extract possible knowledge states of the learners. Every such knowledge state is associated with a set of questions (exercises) which should be answered correctly by a student mastering that knowledge state. In the early work the knowledge states are ordered in an AND-OR-graph, so that the paths through this graph represent possible sequences of knowledge states during the learning process (cf. [Doignon & Falmagne 1985]). These AND-OR-graphs will be

* sponsored by the Deutsche Forschungsgemeinschaft under Grant No. Mo 292/3

described in detail in the next section. Later on Doignon and Falmagne developed a formalism, that accounts for possible inconsistencies in the student´s behaviour, which may arise from lucky guesses and careless errors (cf. [Falmagne & Doignon 1986]).

1.2. The AND-OR-Graph

Specific possible knowledge states are ordered in an AND-OR-graph, such that a surmise relation holds: if a concept represented by an AND-node is mastered all its subordinate concepts will be mastered, too. If a concept represented by an OR-node is mastered, at least one of its subordinate concepts will be mastered, too.

Five exercises from the field of elementary combinatorics and probability are introduced in [Doignon & Falmagne 1985] which should be mastered by students at different levels of competence.

a) Let p be the probability of drawing a red ball in some urn. What is the probability of observing at least one red ball in a random sample of n balls, if the sampling is done with replacement? Give the formula.

b) What is the probability of the joint realization of n independent events, each of which has a probability equal to p?

c) Give the formula for the binominal coefficient $\binom{n}{k}$. Perform the computation for $n = 7$ and $k = 5$.

d) In the experiments of Problem (a), what is the probability of observing exactly k red balls? Give the formula, and perform the computation for $n = 5$ and $k = 3$.

e) Let $P(\mathrm{A})$ be the probability of an event A in a probability space. What is the probability that A is not realized in one trial?

Fig. 1: The sample exercises from [Doignon & Falmagne 1985], p 176

The expected correct results to the exercises are shown in figure 2. Giving the exercises to a group of students may hypothetically lead to observable solution patterns shown in figure 3. The data matrix is ordered after frequency of correct answers to the exercises.

a) $1 - (1 - p)^n$

b) p^n

c) $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

d) $\binom{n}{k} p^k (1 - p)^{n-k}$

e) $1 - p, \; p = P(A)$

Fig. 2: Expected correct results to the sample exercises

	e	c	b	d	a	
1	1	1	1	1	1	5
2	1	1	1	1	1	5
3	1	1	1	0	1	4
4	1	1	1	0	1	4
5	1	1	1	1	0	4
6	1	1	1	1	0	4
7	1	0	1	0	1	3
8	1	0	1	0	1	3
9	1	1	1	0	0	3
10	1	1	1	0	0	3
11	1	1	0	1	0	3
12	1	1	0	1	0	3
13	1	0	1	0	0	2
14	1	0	1	0	0	2
15	1	1	0	0	0	2
16	1	1	0	0	0	2
17	0	1	0	0	0	1
18	0	1	0	0	0	1
19	1	0	0	0	0	1
20	1	0	0	0	0	1
21	0	0	0	0	0	0
22	0	0	0	0	0	0
	18	14	12	6	6	

Fig. 3: Matrix of hypothetical solution patterns to the sample exercises

The first column enumerates the subjects. The next 5 columns denote the answers to the exercises, where "1" means correct answer and "0" no correct answer. The last column gives the number of correct answers per student. The last row shows the total amount of correct answers to every exercise. Figure 4 shows the resulting set of possible knowledge states.

{Ø, {e}, {c}, {e, c}, {e, c, d}, {e, b}, {e, b, c}, {e, b, d, c}, {e, b, a}, {e, b, a, c}, {e, c, d, b, a}}

Fig. 4: Possible knowledge states derived from matrix of fig. 3

Figure 5 shows the Hasse diagram of the knowledge states structured according to the surmise relation. E.g. a student mastering exercise d) is surmised also to master exercises c) and e).

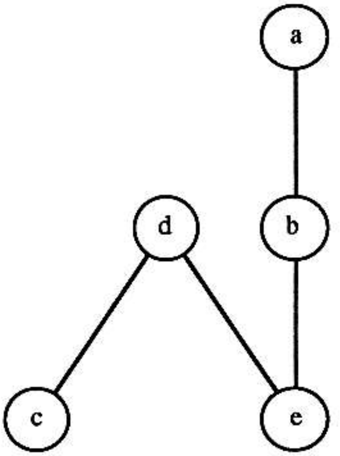


Fig. 5: The Hasse diagram from [Doignon & Falmagne 1985], p 177

Another matrix of solution patterns may be observed together with another solution to exercise a). This solution can be viewed as a special case of exercise d).

$$1 - \binom{n}{0} p^0 (1 - p)^{n - 0}$$

Fig. 6: Another correct solution to exercise a)

The extended data matrix is given in figure 7. The additional entries are marked "23" and "24". They are included between "4" and "5". The extended set of possible knowledge states in figure 8.

		e	c	b	d	a	
	1	1	1	1	1	1	5
	2	1	1	1	1	1	5
	3	1	1	1	0	1	4
	4	1	1	1	0	1	4
new	23	1	1	0	1	1	4
new	24	1	1	0	1	1	4
	5	1	1	1	1	0	4
	6	1	1	1	1	0	4
	7	1	0	1	0	1	3
	8	1	0	1	0	1	3
	9	1	1	1	0	0	3
	10	1	1	1	0	0	3
	11	1	1	0	1	0	3
	12	1	1	0	1	0	3
	13	1	0	1	0	0	2
	14	1	0	1	0	0	2
	15	1	1	0	0	0	2
	16	1	1	0	0	0	2
	17	0	1	0	0	0	1
	18	0	1	0	0	0	1
	19	1	0	0	0	0	1
	20	1	0	0	0	0	1
	21	0	0	0	0	0	0
	22	0	0	0	0	0	0
		20	16	12	8	8	

Fig. 7: Extended matrix from fig. 3

{Ø, {e}, {c}, {e, c}, {e, c, d}, {e, b}, {e, b, c}, {e, b, d, c}, {e, b, a}, {e, b, a, c}, {e, c, d, a}, {e, c, d, b, a}}

Fig. 8: Extended set of knowledge states from fig. 4

Figure 9 shows the AND-OR-graph corresponding to the extended data matrix (cf. fig. 7) and the extended set of possible knowledge states (cf. fig. 8).

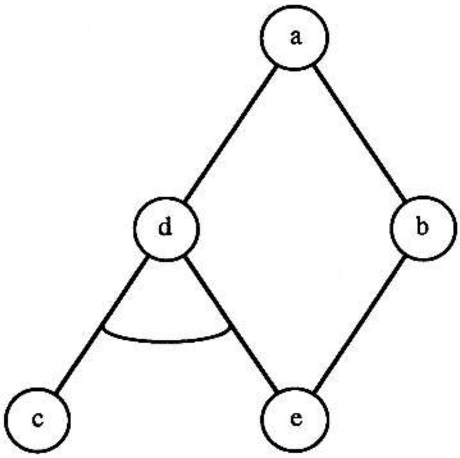


Fig. 9: The AND-OR-graph from [Doignon & Falmagne 1985], p 183

The AND-OR-graph of figure 9 represents the structure of a body of knowledge about elementary combinatorics and probability. The mastering of each exercise requires certain knowledge within this domain. The graph visualizes the surmise relation, that a student mastering an exercise will also master subordinate exercises. A student answering exercise d) correctly is surmised to be able to master both exercises c) and e) as well. This is indicated by the AND-node representing exercise d). Exercise a) is represented by an OR-node. This means from the mastery of exercise a) the mastery of exercise b) or d) will be surmised.

2. The ABSYNT-Diagnostics

The diagnostics in the ABSYNT problem solving monitoring system (cf. [Möbus 1988], [Möbus 1989], [Janke 1989], [Möbus 1990]) consists of a goal-means-relation (cf. [Möbus & Thole, this volume]).It contains many different correct solutions to every of the exercises which will be given to the students during the learning period. ABSYNT includes a visual programming language (cf. [Janke & Kohnert 1988]), which is used by the student to program the presented exercises. The blueprints of the programs, which the students produce during their programming sessions are matched against the correct solutions represented in the diagnostics. The diagnostics is designed to support the student in a way adapted to her/his individual approach to the solution of the current exercise.

2.1. The AND-OR-Graph

The ABSYNT-diagnostics is of a collection of AND-OR-graphs consisting of directed connections and two different sorts of nodes. The round-shaped nodes (cf. fig. 10f) represent possible goals which can be part of a correct solution for the current exercise. The square-shaped nodes represent the ABSYNT-nodes, which are part of the programming language. These are the only terminal nodes. The round-shaped topmost nodes of every graph can be viewed as the goal to compute the entire exercise. An empirical validation of the goals from the goal-means-relation is reported in [Möbus & Thole, this volume].

The exercises by Doignon & Falmagne and the expected correct results of figure 2 can easily be represented in ABSYNT-Diagnostics-AND-OR-graphs. Figures 10a - 10e show a corresponding collection of ABSYNT-Diagnostics-AND-OR-graphs.

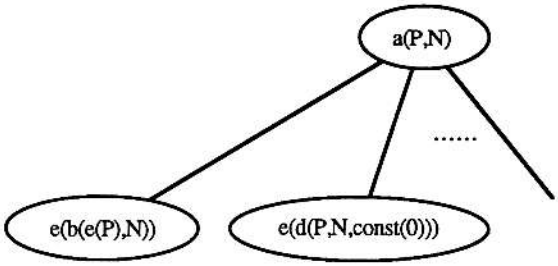


Fig. 10a: ABSYNT-Diagnostics-AND-OR-graph for exercise a)

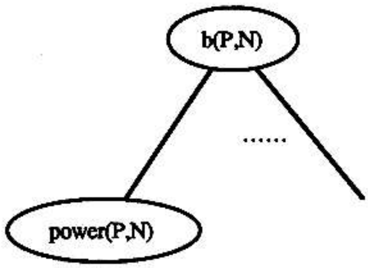


Fig. 10b: ABSYNT-Diagnostics-AND-OR-graph for exercise b)

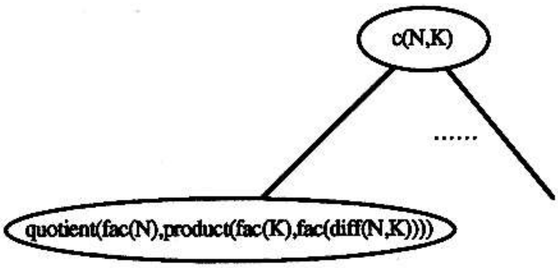


Fig. 10c: ABSYNT-Diagnostics-AND-OR-graph for exercise c)

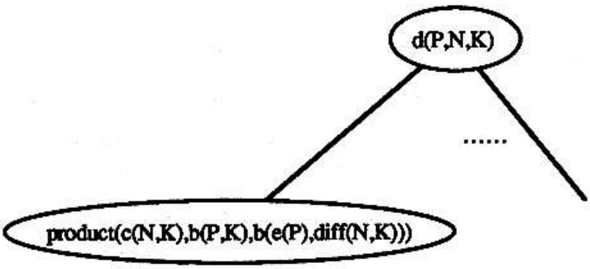


Fig. 10d: ABSYNT-Diagnostics-AND-OR-graph for exercise d)

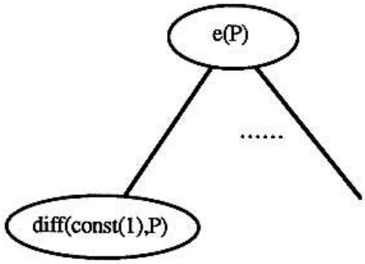


Fig. 10e: ABSYNT-Diagnostics-AND-OR-graph for exercise e)

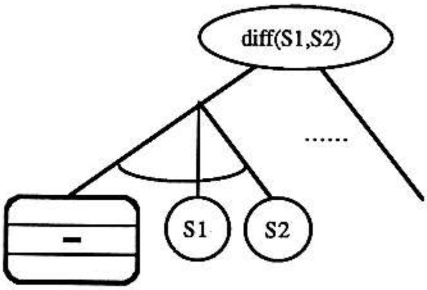
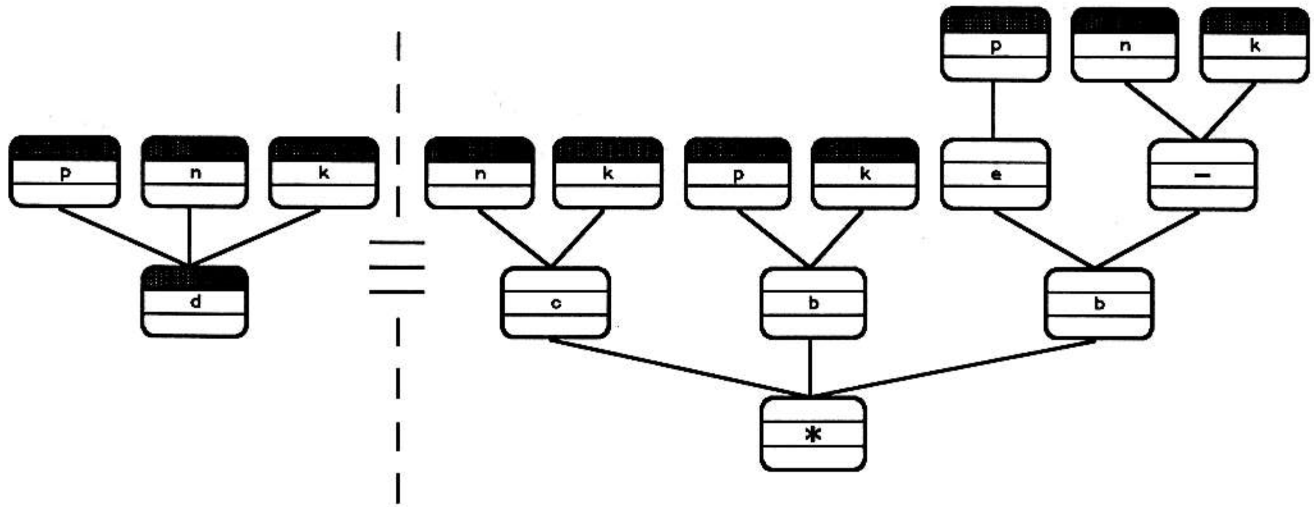


Fig. 10f: ABSYNT-Diagnostics-AND-OR-graph for binary subtraction

The upper-case letters "N", "P", "K", "S1", and "S2" stand for subgoals of the goals-means-relation (cf. [Möbus & Thole, this volume]). The lower-case letters "a", "b", "c", "d", and "e" stand for the respective exercises. "const", "power", "quotient", "product", and "diff" stand for primitives of the ABSYNT programming language, meaning "constant", "power", "division", "multiplication", and "subtraction". Part of the ABSYNT-diagnostics-graph for the subtraction is shown in figure 10f. "fac" stands for "factorial", which is not a primitive; but there exists already a "fac"-graph in the diagnosis, because the factorial is one of our own exercises.

The graphs from figure 10 can be combined to a single AND-OR-graph, which represents the goal inclusions of the exercises following the analysis of Doignon & Falmagne. Figure 11 shows an ABSYNT-frame which is a correct solution for exercise d) programmed with calls to exercises b), c), and e). The corresponding path through the ABSYNT-diagnostics is given in figure 12.



<u>Fig. 11:</u> ABSYNT-frame as solution to exercise d)

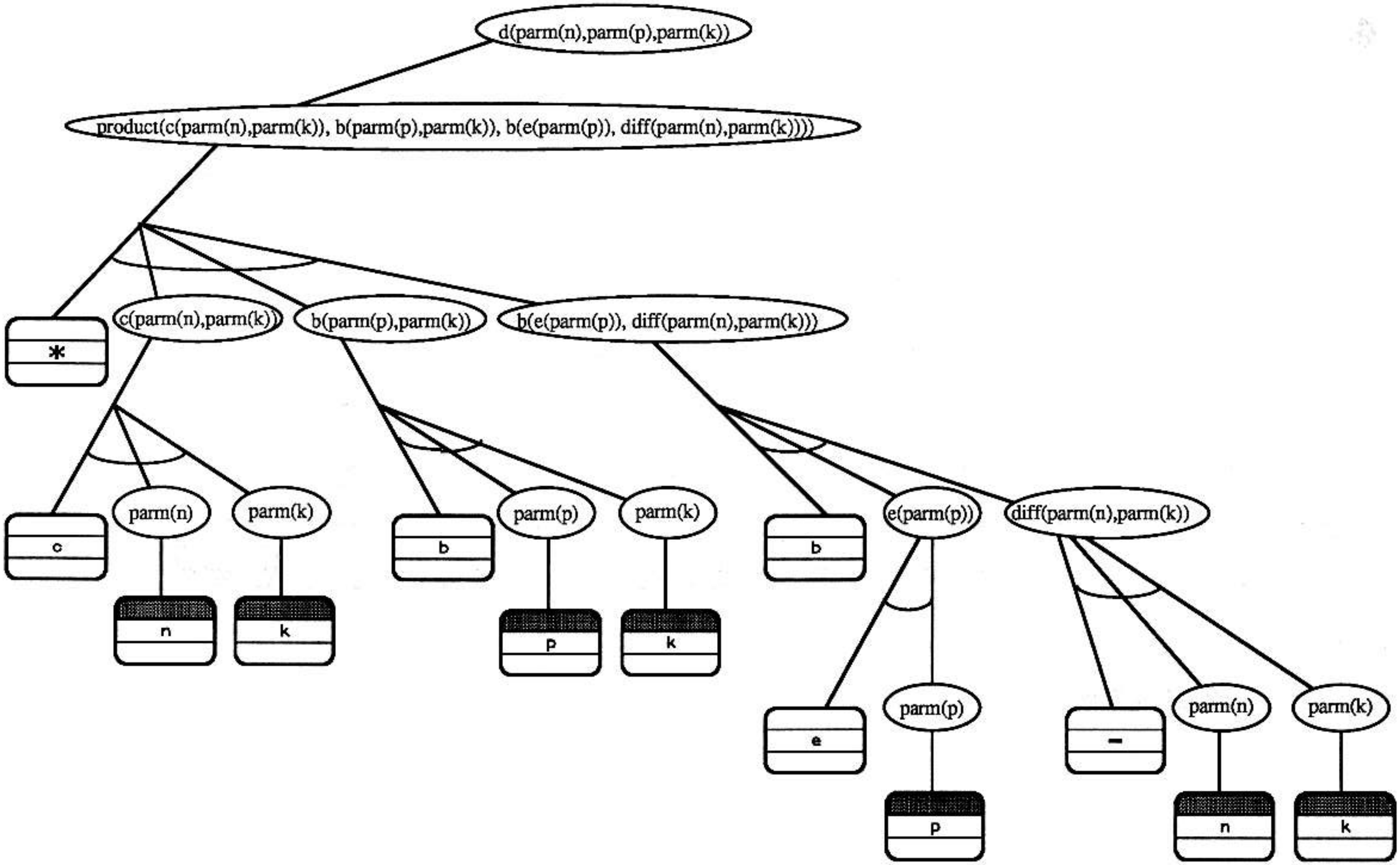


Fig. 12: Path through the ABSYNT-diagnostics for the body of the solution to exercise d) as shown in fig. 11

The sample solution to exercise d), (cf. fig. 11), includes sample solutions to the exercises b), c), and e). This is viewed as a subgoal structure in the ABSYNT-diagnostics. This subgoal structure is shown in figure 13 for the entire set of exercises.

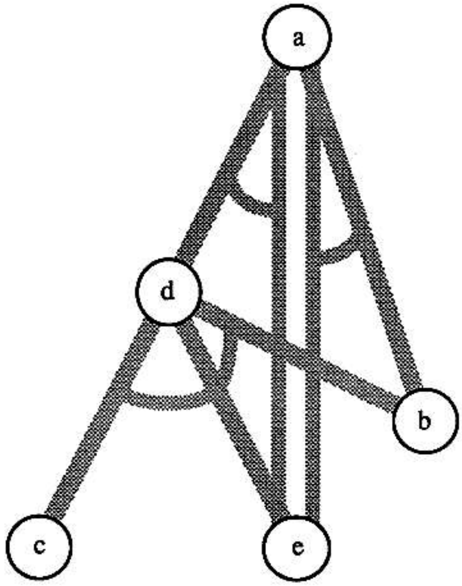


Fig. 13: AND-OR-graph showing subgoal inclusion (ABSYNT-diagnostics)

Figure 14 is a topographically modified reprint of figure 9. It shows the original AND-OR-graph by Doignon & Falmagne. This modification is a necessary for the combined (overlay) representation of figures 13 and 14, which is shown in figure 15 to visually support the comparison.

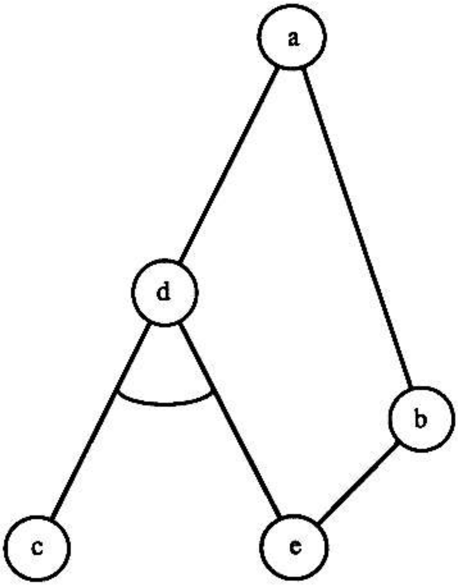


Fig. 14: AND-OR-graph from fig. 9 (surmise relation)

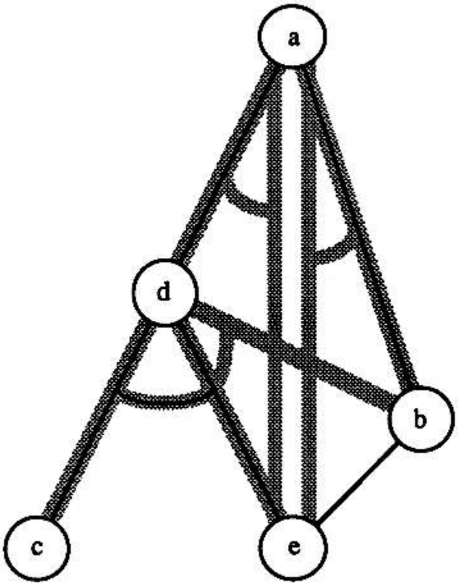


Fig. 15: Combined graphs from fig.s 13 and 14

3. Comparison

3.1. Surmise Relation

The surmise-relation means that from mastery of exercise a) can be inferred the mastery of exercises b) and e) or the mastery of exercises d), c), and e). The nodes of the AND-OR-graph by Doignon & Falmagne represent the exercises; the connections order them according to the surmise relation.

The main purpose of the AND-OR-graph is to serve as a source for the construction of a binary decision tree. Each path through this tree leads to one set of abilities. Each of these sets represents one possible knowledge state. The tree is used to assess the student´s knowledge state, i.e. to select one of the possible knowledge states and to surmise it as the student´s current one, with as few questions (exercises) as possible.

3.2. ABSYNT-Diagnostics

The ABSYNT-AND-OR-graph means that, according to the given problem analysis, exercise a) can be solved with calls to solutions to exercises d) and e) or with calls to solutions to exercises b) and e). Exercise d) can be solved with calls to solutions of exercises b), c), and e). The nodes represent the exercises. The connections relate them after possible abstractions in case of correct solutions.

The AND-OR-graph is used to match the student´s blueprints against the known correct solutions represented in the graph. Depending on the current state of work on the program the known solution with best match is assumed to be the one closest to the student´s intended solution. It will be used to generate the helps offered to the student.

The match is meant to select from the known solutions the one to which the student´s blueprint can be completed. The aid given to the students will be based on this selected solution.

3.2.1. The Missing Connection

In the example the ABSYNT-graph generally contains more connections than the surmise-graph. With one exception the surmise-graph can be said to be more abstract, because it contains only connections which are also part of the ABSYNT-graph.

The interesting exception is the missing connection between b) and e). This connection could easily be included by augmenting the diagnostics with a tautology: The formula of exercise e) applied to itself and again applied to any term always returns the term, because

$$1 - (1 - x) = 1 - 1 + x = x$$

A solution to exercise b) containing a construction like this can be correct; but it is unnecessarily complex, and therefore it is not a good programming style.

3.3. Suitability for Distinct Domains

The surmise-graph reflects the empirical data of observable subjects´ behaviour. The ABSYNT-diagnostics reflects the task structure, which was elaborated in a detailed analysis of the given domain.

The approach of Doignon & Falmagne can be applied to domains with weak task structures, e.g. everyday knowledge. It can also be applied in cases where the subjects´ behavior does not follow the underlying task structure. The ABSYNT-diagnostics can be used to explain individual and task-dependent behaviour of subjects acting according to the task analysis in a well-structured domain. The graphs may be identical or different, depending on the empirical data.

The explaining ability of the ABSYNT-diagnostics can serve as a theoretical foundation for individualized support at a detailed level.

References

Doignon, Jean-Paul; Falmagne, Jean-Claude
Spaces for the Assessment of Knowledge. International Journal of Man-Machine Studies, 23, 1985, pp 175-196

Doignon, Jean-Paul; Falmagne, Jean-Claude
Knowledge Assessment: A Set-theoretic Framework. Internal Report, May 13th, 1986

Janke, Gabriele
Konzeptualisierung eines problemlösezentrierten Hilfesystems. in: STETTER, F.; BRAUER, W. (eds.), Informatik und Schule 1989: Zukunftsperspektiven der Informatik für Schule und Ausbildung, GI-Fachtagung, Munich, November 15th - 17th, 1989, Proceedings; Informatik-Fachberichte 220, Berlin:Springer, 1989

Janke, Gabriele; Kohnert, Klaus
Interface Design of a Visual Programming Language: Evaluating Runnable Specifications. paper presented at the Second Network Seminar of MACINTER held in Berlin/GDR, March 21st - 25th, 1988, in: KLIX, F., STREITZ, N.A., WAERN, N., WANDKE, N. (eds.), Man-Computer Interaction Research MACINTER-II, Amsterdam: Elsevier, 1989, pp 567-581

Möbus, Claus
The Relevance of Computational Models of Knowledge Acquisition for Knowledge Communication with the Problem Solving Monitor ABSYNT. ABSYNT-Report 10/89, July 1989; Lecture given at the "EuroSoar Tutorial and 2nd Workshop" in honor of Allen Newell, Zeegse, Netherlands, June 18th - 22nd, 1989

Möbus, Claus
Plandiagnose bei ABSYNT-Programmen mit attributierten Aufgabengrammatiken. Lecture given at the 2nd workshop "Programmierwissen und seine Modellierung", Department of Psychology, University of Marburg/Lahn, July 25th - 27th, 1988

Möbus, Claus
Toward the Design of Adaptive Instructions and Helps for Knowledge Communication with the Problem Solving Monitor ABSYNT. to appear in: Stepankova, O., Marik, V. (eds.), Proceedings of the CEPES UNESCO Workshop "The Advent of Higher Education", Prague, Oct. 23rd - 25th, 1989, Berlin:Springer, (in press)

Möbus, Claus; Thole, Heinz-Jürgen
Interactive Support for Planning Visual Programs in the Problem Solving Monitor ABSYNT: Giving Feedback to User Hypotheses on the Language Level. This volume.

Schröder, Olaf
A Model of the Acquisition of Rule Knowledge with Visual Helps: The Operational Knowledge for a Functional, Visual Programming Language. This volume.

Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

438

D.H. Norrie H.-W. Six (Eds.)

Computer Assisted Learning

3rd International Conference, ICCAL '90
Hagen, FRG, June 11–13, 1990
Proceedings

Springer-Verlag
New York Berlin Heidelberg London Paris Tokyo Hong Kong