Domain Representation Assisting Cognitive Analysis

Antoaneta Serguieva[1] and **Tariq M Khan**

Abstract. This research examines the application of perception-based reasoning methods [11],[12] to the problem of training. In particular, focus is on how to determine the reason for a learner failing to complete a task. Importantly, the aim is to be able to identify areas of knowledge that are linked to task failure in some specified way, such as missing knowledge and misconceptions. The term cognitive diagnosis is used to refer to the process by which blame for task failure can be assigned, and Zadeh's computational theory of perception is adopted as the basis for the diagnostic capability.

Keywords. computer aided learning, knowledge representation, model based reasoning, generalized constraint based reasoning.

1 INTRODUCTION

Qualitative reasoning is generally accepted as an appropriate method for constructing model-based systems able to perform cognitive diagnosis. Qualitative representation and reasoning capture the way people deal with and manage within the world [2], though these techniques are complimented by quantitative methods where the latter are more appropriate. In [3], for example, a combination of qualitative constraints and numerical reasoning is used to detect physically impossible designs students may produce in thermodynamics. The reason for employing numerical reasoning, as described in [1], is that the calculi underlying qualitative reasoning are relatively weak, and therefore limit its applicability. The computational theory of perception (CTP) complements qualitative representation in providing a further mode of information granulation. While numerical data are singular and qualitative data are c-granular, information is also perceived by people as f-granular (c, crisp as opposed to f, fuzzy). Moreover, CTP is a reasoning method able to process information in all three modes of granulation, including the qualitative mode, through the use of generalized constraints. The contribution of the CTP to the capability of qualitative methods to process and reason with perception based information will allow reformulation of performance analysis.

Research reported here is ongoing and covers areas ranging from representation of domain knowledge using generalized constraints, through corresponding representation of students' knowledge (student modelling) using the notion of perceptions to address the fact that a model of a student is only a representation of their perceptions of the target domain knowledge. This is made possible with the introduction of CTP. The central objective, however, is to develop a diagnostic component that is compatible with the perception based approach to student modelling. In this paper, coverage is limited to the representation of domain knowledge.

[1] Brunel Business School, Brunel University, Uxbridge, United Kingdom email: Antoaneta.Serguieva@brunel.ac.uk, Tariq.Khan@brunel.ac.uk

2 DOMAIN REPRESENTATION

Although there are clearly many ways to categorize knowledge, this research distinguishes between conceptual knowledge and problem-solving knowledge. Several types of problem-solving knowledge are further defined, and sub-types within the more general types. Structural knowledge captures taxonomic and compositional dependencies. Taxonomic knowledge communicates type subtype subordination and inheritance of properties between objects. Compositional relations provide information about the elements of an object. Behavioural knowledge is accounted for by a number of sub-types, where the strength of dependency increases from temporal and co-occurrence, through correlation and enablement, to teleological and causal relations. Suitable techniques are applied to model each type of relation. Diverse modelling techniques have been considered in [9] - from logic and rules, along causal networks and Petri nets, to equations - and then associated with the relational types. It is shown that the association modelling technique relational type is many-to-many rather than one-to-one or all-to-all. Every model stands for a piece of problem-solving knowledge, and a model description is only complete if also indicating the relational type. Furthermore, modifications of each modelling technique are considered to investigate a problem's tolerance toward imprecision, or the benefits from imprecise modelling in terms of tractability, low-cost solution, achieving a solution when no precise information is available, etc. Effectively, this introduces another type of problem solving knowledge - imprecise knowledge. As its sub-types, we adopt the relational types defined in [11],[12]: equal, possibilistic, veristic, probabilistic, probability value, usuality, random set, random fuzzy set, and fuzzy graph. Notably, increasing the depth of imprecision, it is necessary to introduce generalized constraints in problem description, as a generalization of model representation.

As a result, the problem-solving knowledge involving each concept can be described with a set of models arranged in a multi-model space along various modelling dimensions [5],[4], where a model is characterized with its depth in each dimension representing a perspective or level of concept definition. In this research, the following four dimensions are considered: scope, resolution, generality and imprecision. Scope and resolution correspond to the various classes or levels of taxonomic and compositional knowledge, respectively. A higher taxonomy class is consistent with a broader problem scope, and a more detailed compositional level communicates increased resolution. Generality is the dimension along the behavioural types of knowledge, as the stronger relations as causal and teleological tend to rely on more abstract domain principle, while the weaker relations as co-occurrence and temporal involve more specific knowledge like procedures. Imprecision is the dimension exploring the reasonable imprecision in domain representation. The number of relational types involved along this perspective, and therefore its depth, will increase from technical, through industrial and medical, to financial and social domains. Furthermore, the representation along the imprecision perspective requires the use of generalized constraints [11,12] rather than models.

$$\begin{aligned} X \quad & isr_{imprecision} \quad R & (1) \\ r_{imprecision} \quad & \epsilon \quad \{e, d, v, p, \lambda, u, rs, rfs, fg\} \\ e: \quad & equal \\ d: \quad & possibilistic \\ v: \quad & veristic \\ p: \quad & probabilistic \\ \lambda: \quad & probability\ value \\ u: \quad & usuality \\ rs: \quad & random\ set \\ rfs: \quad & random\ fuzzy\ set \\ fg: \quad & fuzzy\ graph \end{aligned}$$

Here, X is a constrained variable, R is a modelled constraining relation, $isr_{imprecision}$ is a variable copula defining the way in which R constrains X, and thus $r_{imprecision}$ is an indexing variable standing for the relational type.

It is necessary to introduce a unifying structuring principle within the domain space. The information along the first three dimensions can be represented using multiple models. The knowledge along the forth dimension can be represented through multiple generalized constraints. Every model can be described as a generalized constraint, while a generalized constraint may not be translated directly or uniquely as a model. Therefore, the introduction of the forth dimension requires the use of multiple generalized constraints as representation elements throughout the domain space. The transformation of the models along the perspectives of scope, resolution and generality, to generalized constraints will be performed to be consistent with the idea that a problem's description is only complete if also indicating the relational type. An analogy with the notation in (1) is kept in definitions (2), (3) and (4).

$$\begin{aligned} X \quad & isr_{generality} \quad R & (2) \\ r_{generality} \quad & \epsilon \quad \{t, co, c, en, te, ca\} \\ t: \quad & temporal \\ co: \quad & cooccurrence \\ c: \quad & correlational \\ en: \quad & enablement \\ te: \quad & teleological \\ ca: \quad & causal \end{aligned}$$

$$X\ isr_{scope}\ R, \quad r_{scope}\ \epsilon\ \{taxonomy\ classes\} \qquad (3)$$

$$X\ isr_{resolution}\ R, \quad r_{resolution}\ \epsilon\ \{compositional\ levels\} \qquad (4)$$

Then we can introduce the description of a generalized constraint in the domain space as

$$\begin{aligned} X \quad & isr_{resolution} r_{scope} r_{generality} r_{imprecision} \quad R & (5) \\ r_{resolution} \quad & \epsilon \quad \{compositional\ levels\} \\ r_{scope} \quad & \epsilon \quad \{taxonomy\ classes\} \\ r_{generality} \quad & \epsilon \quad \{t, co, c, en, te, ca\} \\ r_{imprecision} \quad & \epsilon \quad \{e, d, v, p, \lambda, u, rs, rfs, fg\} \end{aligned}$$

Thus (5) also represents the structuring principles in the multi-perspective multi-constraint domain framework. The representation framework may be instantiated in various target domains.

3 FURTHER RESEARCH

Constructing the domain representation framework is the first step in the overall diagnostic task. An important point is the introduction of the dimension of imprecision. We show in [8],[7] how a domain problem may benefit from exploiting the forth dimension in its representation and solution. The range of domain problems is broaden up in [9]. Future work involves building a complete application in the financial domain, and testing the ideas presented here.

The next step is to develop a corresponding student representation, and to show how it is supported by the domain framework. There will be three principles in exploiting the domain constraints - learner description is *constraint-choice dependent*, *experience related* and *perception based*. Thus, 'hovering over' or 'jetting through' the domain structure and exploring the multifaceted representation of problem-solving knowledge, it is possible to develop a flexible student description. It will be based on multiple generalized constraints and will reflect the student's perception of domain knowledge. Thus the domain framework will act as an explanatory database for the student representation.

Beyond the domain framework and the student description, an important objective is the design of a diagnostic strategy to perform the blame assignment for task failure on specific knowledge. The strategy will involve generalized constraint propagation, which is the reasoning engine in the computational theory of perceptions.

ACKNOWLEDGEMENTS

This work is supported by the EPSRC Grant GR/R51346/01 on cognitive diagnosis in intelligent training (CODIT), and developed within the European Network of Excellence in model based systems and qualitative reasoning (MONET).

REFERENCES

[1] K. de Koning, B. Bredeweg, J. Breuker, and B. Wielinga, 'Model-based reasoning about learner behaviour', *Artificial Intelligence*, **117**, 173–229, (2000).

[2] K. Forbus, 'Using qualitative physics to create articulate educational software', *IEEE Expert*, **12**, 32–41, (1997).

[3] K. Forbus, P. Whalley, J. Everett, L. Ureel, M. Brokowski, J. Baher, and S. Kuehne, 'Cyclepad: An articulate virtual laboratory for engineering thermodynamics', *Artificial Intelligence*, **114**, 297–347, (1999).

[4] T. Khan, K. Brown, and R. Leitch, *Managing Organisational Memory with a Methodology Based on Multiple Domain Models*, 57–76, Proceedings of the Second International Conference on Practical Application of Knowledge Management, 1999.

[5] R. Leitch, M. Chantler, K. Brown, T. Kane, Q. Shen, G. Coghill, A. Steele, M. Ravindranathan, A. Aldea, J. Liddle, S. McDonnchadha, and D. Reay, 'Modelling choices in intelligent systems', *Artificial Intelligence and the Simulation of Behaviour Quarterly*, **93**, 54–60, (1995).

[6] J. Self, 'Model-based cognitive diagnosis', *User Modelling and User Adapted Interaction*, **3**, 89–106, (1993).

[7] A. Serguieva and T. Kalganova, *A Neuro-Fuzzy-Evolutionary Classifier of Low-Risk Investments*, 997–1002, Proceedings of the IEEE International Conference on Fuzzy Systems, IEEE Press , 2002.

[8] A. Serguieva and J. Hunter, 'Fuzzy interval methods in investment risk appraisal', *Fuzzy Sets and Systems*, **142**, 443–466, (2004).

[9] A. Serguieva and T. Khan, *Modelling Techniques for Cognitive Diagnosis*, EPSRC Deliverable Report on Cognitive Diagnosis in Training, Brunel University, 2003.

[10] L. Zadeh, 'Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic', *Fuzzy Sets and Systems*, **90**, 111–127, (1997).

[11] L. Zadeh, *Outline of Computational Theory of Perceptions Based on Computing with Words*, 3–22, Soft Computing and Intelligent Systems, Academic Press, 2000.

[12] L. Zadeh, 'A new direction in AI: Toward a computational theory of perceptions', *Artificial Intelligence Magazine*, **22**, 73–84, (2001).