Pre-publication draft of a paper which appeared in the Proceedings of ASCILITE'88, pp 502-514, Canberra College of Advanced Education.

COGNITIVE DIAGNOSIS USING STUDENT ATTRIBUTIONS

Geoffrey I. Webb

Division of Computing and Mathematics, Deakin University, Victoria, 3217, Australia.

Abstract

This paper details an approach to cognitive diagnosis that enables the inference of detailed models of a student's conceptualisation of a domain. This model is constructed be examining the attributes of the problems that the student has tackled and the student's performance while tackling those problems. A feature network is used to represent educationally relevant domain knowledge.

- This approach has low implementation and operational overheads.
- It provides a detailed model of the student's conceptualisation of the subject domain in terms of elements of knowledge from that domain;
- Student models are not restricted to overlays of predefined correct and/or incorrect knowledge.
- It does not require that the instructional designer anticipate the possible forms of error that may occur.
- It is robust in the face of partial evaluation of student performance.
- It is also robust in the face of the instructional designer's failure to incorporate relevant aspects of the subject domain in the knowledge-base.
- The student models can be executed.
- It supports accurate diagnosis of multiple viewpoints of the domain even when those viewpoints are not anticipated by the instructional designer.
- It can support multiple teaching styles in the one lesson.

Keywords

Student modelling; Cognitive diagnosis

1 Introduction

This paper presents ABM (Attribute-Based Modelling) - a new approach to cognitive modelling. Although developed in the context of student modelling in Computer-Aided Learning, ABM is applicable to any context in which a cognitive model is to be constructed from observations of an external agent's responses to external stimuli.

ABM develops an extremely high level cognitive model. Such a model describes the cognitive system in terms of simple rules that map its inputs directly onto its outputs. This contrasts to extremely low level cognitive models, such as models of neural interactions, and intermediate level cognitive models, such as production system models that posit hypothetical mental operators (Anderson, 1983; Langley, Ohlsson and Sage, 1984; Ohlsson and Langley, 1985).

ABM has been developed as part of the DABIS ICAL project for use with FNBCAL (Feature-Network Based Computer-Aided Learning.) DABIS is described in Webb (1986). FNBCAL is described in Webb (in-press).

ABM is performed in the context of the student examining a body of problems or exercises from the subject domain. Such problems and exercises are hereafter collectively referred to as *instances*.

ABM produces student models by analysing suitable representations of:

- the instances that the student has examined;
- the student's responses when examining those instances; and
- relevant subject domain knowledge.

The relevant subject domain knowledge is represented by a *feature network* (Webb, in-press).

The feature choices in a feature network represent attributes of the instances that the student examines. Thus, feature networks represent the relevant attributes of instances from the subject domain and the inter-relationships between those attributes. This one simple formalism is used to provide a consistent framework which may be used to represent many different aspects of a subject domain. Attributes may represent both instructionally relevant aspects of an instance and/or instructionally relevant aspects of a student's response to that instance.

Currently ABM only supports attributes that form nominal scales. Continuous attributes are not supported.

The student's responses to the examination of an instance are treated as *attributions* to that instance. An attribution to an instance is the ascription of an attribute value to that instance.

2 The analyses

The following list provides an overview of the analyses that ABM produces.

- **Understanding an attribute** - the student has a correct model of the underlying principles represented by the attribute.
- **Non-comprehension of an attribute** - the student has no comprehension (correct or otherwise) of the principles represented by the attribute.
- **Understanding an attribute value** - the student has a correct model of the principles represented by the attribute value.
- **Non-comprehension of an attribute value** - the student has no comprehension (correct or otherwise) of the principles represented by the attribute value.
- **Over-generalisation of an attribute value** - the student over-generalises the principles represented by the attribute value.
- **Under-generalisation of an attribute value** - the student under-generalises the principles represented by the attribute value.
- **Erroneous associations between attribute values** - the student makes an attribution in response to specific combinations of instance attribute values.

2.1 Evaluating a model

A major problem in cognitive diagnosis is evaluating whether a particular model is applicable to a particular agent. The two major forms of evidence available with which to make this diagnosis are the plausibility of the model and whether the agent's behaviour is consistent with the model.

The former factor is accounted for in ABM by preferring analyses higher on the list at the start of Section 2 to those lower on the list.

Unfortunately, the latter form of evidence is reduced in reliability by the likelihood of extraneous factors affecting the agent's behaviour. For example, simple typing errors may make a student behave in a manner quite contrary to her/his true conceptualisation of the subject domain.

As a result, there is a need for evaluation criteria that allow for the possibility of noise. Three types of model need to be accommodated -

1. **invarient output**. Of outputs from the cognitive system of type *A* all are of type *B*. For example, when the student carries a digit from one column to the next s/he always carries the value 1.
2. **varient output**. Of cognitive outputs of type *A* not all are of type *B*. For example, when, the student carries a digit from one column to the next s/he does not always carry the value 1.
3. **random output**. The cognitive outputs of type *A* are random with respect to the range of possible values *B*. For example, when the student carries a digit -from one column to the next s/he randomly carries any digit from 1 to 9.

Three evaluation functions are required:

1. $a\ (a, b)$. A function from pairs of cardinal numbers to true or false. $a\ (a, b)$ (where *a* represents the number of outputs of type *A* that are of type *B* and *b* represents the number of outputs of type *A*) is true if and only if there is sufficient evidence that all outputs of type *A* are of type *B* with due allowance for noise.
2. $b\ (a, b)$. A function from pairs of cardinal numbers to true or false. $b\ (a, b)$ (where *a* represents the number of outputs of type *A* that are of type *B* and *b* represents the number of outputs of type *A*) is true if and only if there is sufficient evidence that not all outputs of type *A* are of type *B* with due allowance for noise.
3. $f\ (X)$. A function from arrays of cardinal numbers to true or false. $f\ (X)$ (where *X* is an array of cardinal numbers indexed by a range of possible outputs *B*, and X_i represents the number of outputs of type *A* that were of type *i*) is true if and only if there is sufficient evidence that outputs of type *A* are not correlated with outputs of type *B*.

In the current implementation of ABM -

- $a\ (a, b) = b > 3\ \&\ \frac{a}{b} \geq .8;$
- $b\ (a, b) = b > 3\ \&\ \frac{a}{b} < .8$; and
- f (X) is evaluated by a chi square test such that $f\ (X)$ is true if and only if $\sum_{i \in B} X_i \geq 8$, where *B* is the range of outputs that indexes *X*, and for $x^2 X$, p <.05.

Note that $\neg a\ (a,b)$ does not necessarily imply $b\ (a, b)$

The current definitions of a and b are not as stringent as may he desired, although they do produce reasonable results. An alternative formulation that is currently under investigation is to use a binomial test to evaluate whether the probability of obtaining the observed distribution of outcomes under a set minimum allowable absolute distribution is high enough to support an alternative to the hypothesis that is represented by the test.

2.2 Notation

Let C_i denote the correct attribute values of the instance *i*. For an examination, *e*, of an instance, *i* -

- A_e denotes the set of attribute values that the student has (correctly or incorrectly) identified as those of the instance *i* during exercise *e*;

- U_e denotes the set of attributes that applied to the instance *i* for which the student had the opportunity to but did not identify any attribute values (correct or otherwise); and
- N_e denotes the set of attributes for which the student did not have an opportunity to make an attribution.

For any examination, *e*, the instance examined will be denoted by I_e. The set of all examinations of an instance *i* will be denoted by E_i. K_e denotes the set of correct attribute values for the instance examined in examination *e*.
(This is introduced solely for notational convenience as $K_e = C_{I_e}$). R_a denotes the attribute to which an attribute value *a* belongs.

Let *#S* denote the number of elements in a set *S*.

2.3 Understood attributes and attribute values

An essential type of analysis is whether an agent correctly conceptualises an aspect of the subject domain. Cast in terms of attribute values and attributions, if an agent understands an attribute value s/he can be expected to apply it correctly. That is, s/he can be expected to invariably attribute that attribute value to all and only instances of which it is an attribute value. The understanding of an attribute value can be inferred if such behaviour is observed.

For example, analysing an arithmetic problem in terms of its constituent columns, one attribute of the examination of a particular column may represent whether or not a digit should be carried to the next column. This may be represented by an attribute with the two values **Carry** and **No Carry** representing respectively whether or not a digit should be carried from the current column. It can be inferred that **Carry** is understood if the student identifies it as applying to all and only columns to which it does apply.

For any attribute value, a, infer $understood_a$

if

$$\alpha(\#\{i : a \in A_i \ \& \ a \in K_i\}, \#\{i: a \in K_i \ \& \ R_a \notin N_i\})$$

and

$$\alpha(\#\{i : a \in A_i \ \& \ a \in K_i\}, \#\{i : a \in A_i\}).$$

In addition to evaluating whether a student understands individual attribute values it is also desirable to be able to evaluate whether s/he understands an attribute as a whole. It can be inferred that an attribute is understood if all of the attribute values in it are understood.

Infer $Understood_r$ is true for an attribute *r*

if

$$\forall a(R_a = r \Rightarrow understood_a)$$

There are two educational implications that follow from an inference that a student understands an attribute or attribute value. First, as the student understands the relevant aspect of the domain there is no need to continue examining it beyond any requirements for reinforcement and maintenance. Second, any material for which understanding the attribute or attribute value is a prerequisite may be examined.

2.4 The non-comprehension of an attribute or attribute value

There are many ways in which an agent may fail to fully understand an aspect of the domain. These fall into two major classes. Either the agent may not understand the concept but none the less have a consistent incorrect conceptualisation in its place or s/he may have no fixed conceptualisation whatsoever. Non-comprehension represents the second of these classes of failures in understanding.

If a student has no fixed conceptualisation of an aspect of a domain then s/he can be expected to either not attempt to tackle that aspect of the domain or to make random attributions when s/he does. (Note that if s/he adopts a fixed incorrect strategy then s/he will fail into the former rather than the latter class of failure of understanding.)

As a result, it can be inferred that an agent has a non-comprehension of an attribute or attribute value if s/he either behaves at random with regard to it or refuses to make attributions with regard to it.

For any attribute, r, $non\text{-}comprehended_r$

if

$f(M)$, where M is a two dimensional matrix with the indices a and b such that

$$\forall a(R_a = r)$$

$$\forall b(R_b = r)$$

and

$$M_{ab} = \#\{i : a \in A_i \,\&\, b \in K_i\} + \frac{\#\{i : \exists r(R_r = R_a \,\&\, r \in U_i)\}}{\#\{i : i \in R_{a\}}}$$

Note that the determination of the values for the cells of this matrix is greatly complicated by the need to account for occasions on which the student declines to identify an attribute value from an attribute that applies to an instance. In this case the above equation ensures that the student is taken (for the purposes of this analysis) to have equally partially identified each attribute value for that attribute.

For any attribute value a, $non\text{-}comprehended_a$

if

$f(M)$, where M is a one dimensional matrix with the index b such that

$$\forall b(R_b = R_a),$$

and

$$M_b = \#\{i : a \in A_i \,\&\, b \in K_i\} + \frac{\#\{i : \exists r(R_r = R_a \,\&\, r \in U_i)\}}{\#\{i : i \in R_{a\}}}$$

Similarly to the case with non-comprehension of an attribute, this analysis is greatly complicated by the need to account for failures to identify an attribute value for an attribute.

Inferring that a student has a non-comprehension of an attribute or attribute value indicates that there is a need to teach the student about the principles under lying that attribute or attribute value and how to apply those principles to instances from the domain.

2.5 Over and Under-generalisations

A frequent form of erroneous conceptualisation of an aspect of a domain is to correctly master a principle except for either over or under-generalising it to instances in which it is not appropriate.

Over-generalisation occurs when an agent correctly applies a principle when it ought to be applied, but is also inclined to apply it when it is not appropriate to do so.

For any attribute value a, infer $\textit{over-genera1ised}_a$

if

$$a(\#\{i : a \in A_i \ \&\ a \in K_i\}, \#\{i : a \in K_i \ \&\ a \notin N_i\})$$

and

$$b(\#\{i : a \in A_i \ \&\ a \in K_i\}, \#\{i : a \in A_i\})$$

Under-generalisation occurs when an agent only applies a principle when it ought to be applied but fails to apply it to all instances to which it applies.

For any attribute value, a, infer $\textit{under-genera1ised}_a$

if

$$a(\#\{i : a \in A_i \ \&\ a \in K_i\}, \#\{i : a \in A_i\})$$

and

$$b(\#\{i : a \in A_i \ \&\ a \in K_i\}, \#\{i : a \in K_i \ \&\ a \notin N_i\})$$

Under and over-generalisations are frequently not random. Usually there will be an underlying conceptual error that guides the process. An insight into the nature of such an error is provided by the direction of an under or over-generalisation.

An attribute value is held to be over-generalised toward another attribute value if the over-generalisations of the former attribute value invariably take the form of attributing the former attribute value to an instance which exhibits the latter attribute value.

For any pair of attribute values a and b, infer $\textit{over-generalised toward}_{ab}$

if

$$\textit{over-genera1ised}_a$$

and

$$a(\#\{i : a \in A_i \ \&\ b \in K_i\}, \#\{i : a \in A_i \ \&\ a \notin K_i\})$$

An attribute value is held to be under-generalised toward another attribute value if the under-generalisations of the former attribute value invariably take the form of attributing the latter attribute value to an instance which exhibits the former attribute value.

For any pair of attribute values a and b, infer $\textit{under-generalised toward}_{ab}$

if

$$\textit{under-genera1ised}_a$$

and

$$a(\#\{i : b \in A_i \ \&\ b \in K_i\}, \#\{i : a \in K_i \ \&\ a \notin A_i\}).$$

2.6 Erroneous Associations

Erroneous associations represent more complex forms of mis-conceptualisation to those described above. A frequent form of mis-conceptualisation is to believe that relationships hold between aspects of the domain that do not actually hold. An erroneous association represents exactly this form of mis-conceptualisation.

An association holds between an attribute value and a set of attribute values with which it is associated if the presence of the set of associated attribute values invariably prompts the student to attribute the given attribute value to an instance. An association is erroneous if it does riot hold in the body of instances that the student has examined.

$Associated_{X_a}$ is true for any set of attribute values X and attribute value a

if

$$\forall Y(X \subset Y \Rightarrow \neg\exists b(b \neq a \;\&\; Associated'_{Yb} \;\&\neg Associated'_{Ya}),$$

$$\forall b(b \in X \Rightarrow \exists i(a \in K_i \;\&\; b \notin K_i))$$

and

$$a(\#\{i : a \in A_i \;\&\; X \subseteq K_i\}, \#\{i : a \in A_i\}).$$

where

$Associated'_{Xa}$ is true for any set of attribute values X and attribute value a

if

$$a(\#\{i : a \in A_i \;\&\; X \subseteq K_i\}, \#\{i : a \in A_i\}).$$

The first of these clauses restricts association sets to the minimal sets that describe the association. If a subset of a set of attribute values is associated with an attribution then the superset is not treated as the associated set. Rather, the smaller set is preferred as a more economical explanation of the student's conceptualisation of the attribute value.

The second clause excludes from X attribute values that are invariably present when a is present.

The final clause enables X to be any other set of attribute values for which there is sufficient evidence that all are present whenever the student attributes a to an instance.

A cognitive model solely containing associations is equivalent to a production system model of the cognitive system. Representing that the presence of a set of attribute values A is associated with an attribution b is equivalent to forming a production rule with the presence of the set of attribute values A as the antecedent and the attribution of b as the consequent.

An understanding is equivalent to a (non-erroneous) association between the presence of an attribute value and its attribution to an instance. Thus, a cognitive model in terms of understandings and erroneous associations alone is equivalent to a production system based cognitive model.

Non-comprehensions and under and over-generalisations extend beyond the representational power of a production system based cognitive model.

If the system's knowledge-base contains all instance attributes of which a student takes account in making attributions then all under and over-generalisations will be further specified by erroneous associations that detail the conditions under which the under or over-generalisations take place. Even so, under and over-generalisations provide extremely valuable diagnostic information in that they indicate that the erroneous associations represent relatively minor bugs in a basically correct conceptualisation of the domain knowledge.

However, they are also extremely valuable when the instructional designer has failed to anticipate all relevant attributes from the subject domain. In this case, they still provide useful diagnostic information even if it is not possible to detect erroneous associations that further specify the nature of the mis-conceptualisation.

3 General considerations

The models that ABM produces are executable. That is, given an ABM model and a novel instance it will be possible to predict the attributions that the student will make to that instance. This is an important feature as it enables an intelligent instructor to create of plans that incorporate student actions as well as the tutor's.

Another beneficial feature of ABM is its low computational overheads. Unlike most alternative approaches to deriving student models from behavioural observations, the computational overheads of producing an ABM model are negligible. As a consequence, it is possible to construct and update a detailed student model in real time during a lesson, and to use that model for lesson management.

ABM models should be sharply distinguished from simple overlay models (Carr and Goldstein, 1977). An overlay model assigns to the student a subset of the system's domain knowledge. An ABM model containing solely understandings is equivalent to an overlay model. Non-comprehensions, over and under-generalisations and erroneous associations all extend the descriptive power of an. ABM model beyond that of an overlay model.

The ABM approach to cognitive modeling should also be distinguished from the use of buggy libraries (Anderson., Boyle and Reiser, 1985; Brown and Burton, 1978). A buggy library is a collection of possible knowledge errors that may be ascribed to the student. ABM does not require the instructional designer to anticipate the possible forms of student error that may occur, as is the case with buggy library approaches to student modelling. Rather, the instructional designer need only anticipate the aspects of the subject domain that may be relevant to diagnosing the student's mis-comprehensions. The diagnostic system automatically determines the exact nature of any particular bug without reference to a library of possible bugs.

Further, even if the instructional designer fails to anticipate the relevance of an aspect of the domain, the system will still be able to make accurate diagnoses (in terms of understandings, non-comprehensions, and over and under-generalisations) even though those diagnoses will not be as detailed as would otherwise have been possible (that is, will not include the relevant erroneous associations.)

A problem faced by many instructional systems is the inability to handle multiple viewpoints of the subject domain (Wenger, 1987). That is, even though most domains can be approached from many different perspectives (for example, by using many different sets of operators on problems in the domain), most tutoring systems can only accommodate a single viewpoint of the subject matter. As a result, student's that have a correct but different understanding of the domain to the tutoring system's will be diagnosed and treated as having an incorrect conceptualisation thereof. ABM does not suffer from this deficiency so long as the attributes of instances and of the student's performance that are dealt with are non-subjective observable details. Under this condition, ABM can provide accurate models that are viewpoint independent.

A related point is that ABM is not tied to a particular approach to teaching (other than that a lesson must occur in the context of examining instances from a domain.) The one student model may be used with substantially different teaching styles allowing lessons to select a teaching style that best suits the learning style of the individual learner.

An implementation of ABM has been used with great success in a simple DABIS (Webb, 1986) lesson on English word classes (see Webb, forthcoming). Work is now in progress to create a lesson on a simple procedural skill, unification, with the goal of demonstrating that the approach is applicable to procedural as well as analytical domains.

References

[1] Anderson, J.R. (1983). *The Architecture of Cognition*. Harvard University Press, Cambridge, Mass.

[2] Anderson, J. R., Boyle, C. F. and Reiser, B. J. (1985). **Intelligent tutoring systems**. *Science*, 228:456 - 462.

[3] Brown, J. S. and Burton, R.R. (1978). **Diagnostic models for procedural bugs in basic mathematical skills**. *Cognitive Science*, 2:155 - 192.

[4] Carr, B. and Goldstein, I. P. (1977). **Overlays: A Theory of Modelling for Computer-Aided Instruction.** *AI Lab Memo 406*, Massachusetts Institute of Technology, Cambridge, Mass.

[5] Langley, P., Ohlsson, S. and Sage, S. (1984). **A machine Learning Approach to Student Modeling**. *Tech. Rep. CMU-RT-TR-84-7*, The Robotics Institute, Carnegie-Mellon University.

[6] Ohlsson, S. and Langley, P. (1985). **Identifying Solution Paths in Cognitive Diagnosis**. *Tech. Rep. CMU-RI-TR-85-2*, The Robotic Institute, Carnegie-Mellon University.

[7] Webb, G. I. (1986). **Knowledge Representation in Computer-Aided Learning: The Theory and Practice of Knowledge-Based Student Evaluation and Flow of Control**. *PhD thesis*, La Trobe University School of Mathematical and Information Sciences.

[8] Webb, G. I. (forthcoming). **Attribute-based cognitive diagnosis**. To be published.

[9] Webb, G. I. (in-press). **A knowledge-based approach to computer-aided learning**. To be published in the *International Journal of Man-Machine Studies*.

[10] Wenger, E. (1987). *Artificial Intelligence and Tutoring Systems*. Morgan Kaufmann, Los Altos, CA.