Formalising the knowledge content of case memory systems

A D Griffiths and D G Bridge
Department of Computer Science, University of York, YORK YO1 5DD, UK
{tony|dgb}@minster.york.ac.uk

December 22, 1994

Abstract

Discussions of case-based reasoning often reflect an implicit assumption that a case memory system will become better informed, i.e. *will increase in knowledge,* as more cases are added to the case-base. This paper considers formalisations of this 'knowledge content' which are a necessary preliminary to more rigourous analysis of the performance of case-based reasoning systems. In particular we are interested in modelling the learning aspects of case-based reasoning in order to study how the performance of a case-based reasoning system changes as it accumlates problem-solving experience. The current paper presents a 'case-base semantics' which generalises recent formalisations of case-based classification. Within this framework, the paper explores various issues in assuring that these sematics are well-defined, and illustrates how the knowledge content of the case memory system can be seen to reside in both the chosen similarity measure and in the cases of the case-base.

1 Introduction

Responsible engineering of knowledge-based systems should involve setting error bounds on the performance of the system and, in the case of systems that are to 'learn from experience', guaranteeing that the accumulated experience will actually enhance the performance of the system. We believe that a necessary preliminary to these activities is to be able to state formally the *knowledge content* of a knowledge-based system. Given that a system has 'knowledge' to guide its operation and that it may 'gain knowledge' through problem-solving experience, of what does this knowledge consist and how may it be compared between systems? A formalisation suitable for some aspects of the operation of a case memory system is discussed here.

Some theoreticians [7] [8] have attempted to describe knowledge content from a *logical* point of view, i.e. in terms of the logical propositions that might be derived from a case memory system. These presentations necessarily use non-monotonic logics in order to capture the way in which the set of propositions entailed by a case-base changes as new cases are added. This approach has not however had much success, we believe, in explicating knowledge content and has no demonstrated application to the engineering issues which we believe must be addressed. In contrast the view adopted here is a *functional* one in that the knowledge content of a case memory is modelled as a mapping between input and output domains.

A functional viewpoint has been used in the work of Dearden and Harrison [3] in the engineering of the user interface of case memory systems and, of more relevance to our own work, it has also been assumed in much of 'computational learning theory'. This theoretical study of learning systems models the current state of a learner as a function which is adjusted to accommodate the data presented to the system. We are currently making progress in the study of various case-based learning algorithms within this framework [4].

2 Generalised Case-Base Semantics

We consider the general case where the case memory system is intended to respresent a mapping between an input domain X_1 and an output domain X_2. Additionally, X_1 and X_2 are both finite sets. Using the terminology of Dearden & Harrison [3] we refer to the elements of X_1 & X_2 as 'descriptions' and 'reports' respectively. In [3], the knowledge content of the case memory system is embodied in the 'retrieve' function, which maps from a domain of problem statements into the space of partial orders over the cases in the case-base. This gives a very general view of the operation of a case memory system and one that is clearly necessary for discussing the interface properties of such systems. We opt here for a simpler model; we assume that for each description $x_1 \in X_1$ there is a unique, 'correct' report $x_2 \in X_2$ which we would like the case-memory system to return, so that the ideal behaviour of the system is described by a functional mapping from the set $(X_1 \rightarrow X_2)$. This assumption is common in the non-Bayesian forms of learning theory, and in that context the mapping is often called a 'target function'. Here, the notion that a single output value is appropriate to describe the operation of the system is something of a restriction, but has the advantage that it allows a straightforward definition of the error in a system's output, and allows the role of the similarity measure in the system to be demonstrated more clearly.

Following the work of Jantke [5] a case memory system is modelled as a pair $\langle CB, \sigma \rangle$ consisting of a case-base CB and a similarity measure σ. The case-base is a set of description-report pairs. i.e. CB is an object of type:

$$CB : \mathcal{P}(X_1 \times X_2)$$

In the current paper, we assume that cases in the case-base are free from observational error, so that given some target function $f \in (X_1 \rightarrow X_2)$:

$$\forall x_1 \in X_1, x_2 \in X_2 \cdot (x_1, x_2) \in CB \rightarrow f(x_1) = x_2 \qquad (1)$$

This is equivalent to the statement $CB \subseteq f$, and so CB can also be seen as a partial function between X_1 and X_2. The task of a case memory system is to interpolate the *partial* function represented by the case-base so that the system as a whole represents a *total* function which will return some report value for any description presented to the system. The two-place model $\langle CB, \sigma \rangle$ emphasises that the interpolation made by the system is *dependent on the choice of similarity measure.* For the purposes of the model we consider a similarity measure to be a function over pairs of descriptions returning a normalised real value indicating the degree of similarity between the two objects:

$$\sigma : (X_1 \times X_1) \rightarrow [0, 1]$$

Given a problem description $x \in X_1$ we can therefore define the set of nearest neighbours of x with respect to CB and σ:

$$\forall x \in X_1 \cdot NN(x, CB, \sigma) \mathrel{\hat{=}} \{(x_1, x_2) | (x_1, x_2) \in CB \wedge \forall (x_1', x_2') \in CB \cdot \sigma(x, x_1) \geq \sigma(x, x_1')\} \qquad (2)$$

The essence of case-based reasoning is very simple. We wish the case memory system to return the report from the case whose description is most similar according to the similarity measure σ to the description of the current problem. In order to consider $\langle CB, \sigma \rangle$ as a functional mapping between X_1 and X_2 we need only resolve the question of 'ties'. That is, where the set of nearest neighbours defined in equation (2) contains more than one exemplar from the case-base, we must specify a way of choosing the case used to justify the system's output. Jantke [5] offers three suggestions for resolving such ties:

1. In the case of ties, the output of the system is undefined. That is, the case memory system is interpreted as a partial function.

2. Ties are resolved by a preference ordering over *descriptions*.

3. Ties are resolved by a preference ordering over *reports*.

Option 1 is an undesirable state of affairs in that we may often have perfectly good grounds for choosing one 'equally similar' case over another. In the latter two options the system is equipped with some additional domain knowledge which allows the nearest neighbour with the most preferred description or report to dictate the output. (This might be interpreted, for example as knowledge of the *a priori* plausibility of different descriptions or reports.) In what follows, we consider a general model of preference relations over the set of reports X_2. Option 2 is passed over for the moment principally because this knowledge is potentially subsumed in the similarity measure. Thus we define the function $f_{\langle CB,\sigma\rangle}$ represented by a case-memory system $\langle CB,\sigma\rangle$ as (c.f. [5, p.219]):

$$f_{\langle CB,\sigma\rangle}(x_1) = x_2 \text{ where } \max_{\sqsupseteq}\{x_2' | (x_1', x_2') \,\epsilon\, NN(x_1, CB, \sigma)\} = \{x_2\} \tag{3}$$

$$f_{\langle CB,\sigma\rangle}(x_1) = x_2 \text{ where } NN(x_1, CB, \sigma) = \{\} \wedge x_2 \,\epsilon \max_{\sqsupseteq} X_2 \tag{4}$$

$\sqsupseteq$ is a partial order defining preferences over X_2 so that for a pair of reports $x, x' \,\epsilon\, X_2$, $x \sqsupseteq x'$ reads 'x is preferred to x'', and $\max_{\sqsupseteq}$ are the maxima with respect to that ordering defined as follows:

$$\max_{\sqsupseteq} X = \{x \,\epsilon\, X | \forall x' \,\epsilon\, X \cdot x' \sqsupseteq x \rightarrow x = x'\} \tag{5}$$

The 'case-base semantics' given in equations (3) & (4) are a generalisation of interpretations given elsewhere for the special case of case-based systems for classification. This is illustrated below.

Example 1 Case-base semantics for classification

If we take the special case where the output domain $X_2 = \{0, 1\}$, then we are concerned with the task of classification. Given a problem description from X_1, we require that our system outputs a 'yes' or 'no' result classifying the problem description as an instance of some concept. Elsewhere in our work [4] we have used the following equation as semantics for a case-based classifier, related to the 'standard semantics' of Jantke and Lange [6, p.142] (see also [9, p.84]).

$$f_{\langle CB,\sigma\rangle}(x) = \begin{cases} 1 & if \;\; \exists (x_{pos}, 1) \,\epsilon\, CB \cdot \forall (x_{neg}, 0) \,\epsilon\, CB \cdot \sigma(x, x_{pos}) > \sigma(x, x_{neg}) \\ 0 & otherwise \end{cases} \tag{6}$$

Informally, a point x from X_1 is positively classified by $h_{\langle CB,\sigma\rangle}$ if and only if there is a stored positive exemplar x_{pos} which is strictly more similar to x according to the chosen similarity measure σ than any of the stored negative exemplars x_{neg}. In other words this interpretation resolves 'ties' between equally similar near neighbours by preferring '0' reports to '1' reports; this might be called a conservative classification strategy. Formally, equation (6) is an instantiation of equations (3) & (4) given the preference order $\sqsupseteq = \{(0, 0), (0, 1), (1, 1)\}$. □

Clearly equations (3) & (4) require a little more discussion in the general case where $\sqsupseteq$ is any partial order. For an empty case-base equation (4) is under-constrained and allows a range of possibilities. In the case of Example 1, '0' is (trivially) preferred to every other report. This value will therefore be returned by equation (6) on any description for an empty case-base. Any such solution is admitted by equation (4) as long as for any description, the value returned is a maximum for the entire set of reports X_2. For a non-empty case-base however, there may sometimes be more than one maximum defined on the set of reports of nearest neighbours with respect to some point in the input domain X_1. In this case there is no functional interpretation for $\langle CB, \sigma \rangle$ and we consider that $f_{\langle CB,\sigma\rangle}$ is undefined. Instead, we constrain σ and $\sqsupseteq$ so that $f_{\langle CB,\sigma\rangle}$ always returns a single value. Specifically, for any domain value $x \in X_1$, the set of nearest neighbours with respect to σ must have an upper universal bound which is maximal with respect to the preference ordering $\sqsupseteq$ and in addition is comparable to all other elements in the nearest set. These constraints are developed in the following results:

Definition 2 Upper Universal Bound $x \in X$ *is an upper universal bound for a poset* $\langle X, \sqsupseteq \rangle$ *iff* x *is a member of the maxima of* $\langle X, \sqsupseteq \rangle$ *and also* x *is comparable to every element* $x' \in X$ $(x \in \max_{\sqsupseteq} X \wedge \forall x' \in X \cdot x \sqsupseteq x')$.

Lemma 3 *Given a poset* $\langle X, \sqsupseteq \rangle$, $\sqsupseteq$ *defines a single maximum* x_{max} *on* X $(|\max_{\sqsupseteq} X| = 1)$, *iff* x_{max} *is an upper universal bound for* X.

Proof: **a) Only if.** Assume there is one value x_{max} in the maxima. It must be shown that $x_{max} \sqsupseteq x'$ for any $x' \in X$, hence x_{max} is an upper universal bound to X (Definition 2). From equation (5):

$$\forall x' \in X \cdot x' \sqsupseteq x_{max} \rightarrow x_{max} = x' \tag{7}$$

Additionally, there is only one value x_{max} satisfying this equation. Hence:

$$\forall x' \in X \cdot x' \neq x_{max} \rightarrow \exists x'' \in X \cdot x'' \neq x' \wedge x'' \sqsupseteq x' \tag{8}$$

For any $x' \in X$, by equation (8) either $x_{max} = x'$ and hence $x_{max} \sqsupseteq x'$ by reflexivity, or there is some x'' which is distinct from x' and which is preferred to x' in the preference ordering. Equally, equation (8) applies to x'', and either $x'' = x_{max}$ or there is some x''' preferred to x''. Since $\sqsupseteq$ is a partial order and defines an acyclic graph on X, then eventually the path through the graph of the preference ordering must reach x_{max} for finite sets X. Hence by transitivity $x_{max} \sqsupseteq x'$ for any $x' \in X$ and x_{max} is an upper universal bound for X.

b) If. Assume $\sqsupseteq$ defines an upper universal bound for X, i.e. for any $x_{max} \in \max_{\sqsupseteq} X$ then $x_{max} \sqsupseteq x'$ for any $x' \in X$. Since $\sqsupseteq$ is a partial order and anti-symmetric, there can only be one such x_{max}. □

Definition 4 Admissible Preference Relation *A preference relation* $\sqsupseteq$ *is admissible with respect to a similarity measure* σ *and a function space* F, *iff for every domain value* $x \in X_1$ *and for every non-empty case-base valid for some* $f \in F$ *(i.e.* $CB \subseteq f$*),* $\sqsupseteq$ *defines an upper universal bound for the set* $\{x_2 | (x_1, x_2) \in NN(x, CB, \sigma)\}$.

Corollary 5 *The generalised case-base semantics expressed in equation (3) will be defined, i.e. the preference relation will choose only a single value from the set of nearest neighbours, iff* $\sqsupseteq$ *is an admissible preference relation with respect to* σ *and* F.

Since both the preference relation and the similarity measure share responsibility for discriminating between more and less applicable cases, we can picture a 'trade-off' between the information content of the similarity measure and the preference relation. That is, a less informed preference relation is compensated for by a more discriminating similarity measure and vice versa. In the light of this, the following proposition considers the case where $\sqsupseteq$ contains no extra information, and shows the minimum constraint on the similarity measure entailed by Definition 4 in this situation.

Proposition 6 $\sqsupseteq_I$, *the identity relation on* X_2, *is an admissible preference relation with respect to a measure* σ *and a space of functions* F *iff for any function* $f \in F$ *and for any domain values* $x, x', x'' \in X_1$ *then* f *maps to a different value on* x' *and* x'' *only if the similarities* $\sigma(x, x')$ *and* $\sigma(x, x'')$ *return different values. i.e* $\sqsupseteq_I$ *is an admissible preference relation with respect to* σ *and* F *iff*

$$\forall f \in F \cdot \forall x, x', x'' \in X_1 \cdot f(x') \neq f(x'') \rightarrow \sigma(x, x') \neq \sigma(x, x'') \tag{9}$$

Proof: Note that $x \,\epsilon\, X \to x \,\epsilon\, \max_{\sqsupseteq_I} X$, hence $\max_{\sqsupseteq_I} X$ is a singleton (entailing that $\sqsupseteq_I$ defines an upper universal bound on X by Lemma 3) iff $|X| = 1$. **a) If.** Therefore it must be shown that if equation (9) holds for some σ and F then for any domain function $f \,\epsilon\, F$, for any non-empty case-base $CB \subseteq f$ and for any problem instance $x \,\epsilon\, X_1$, the nearest neighbours in CB to x under the measure σ share a common report x_2. For any such x & x_2, from equation (2):

$$\exists x_1 \,\epsilon\, X_1 \cdot ((x_1, x_2) \,\epsilon\, CB \wedge \forall x_1' \,\epsilon\, X_1, x_2' \,\epsilon\, X_2 \cdot (x_1', x_2') \,\epsilon\, CB \to \sigma(x, x_1) \geq \sigma(x, x_1')) \tag{10}$$

It must be shown that x_2 is unique. i.e.

$$\forall x_2' \,\epsilon\, X_2 \cdot x_2' \neq x_2 \to \forall x_1' \,\epsilon\, X_1 \cdot ((x_1', x_2') \not\epsilon\, CB \vee \exists x_1'' \,\epsilon\, X_1, x_2'' \,\epsilon\, X_2 \cdot (x_1'', x_2'') \,\epsilon\, CB \wedge \sigma(x, x_1') < \sigma(x, x_1'')) \tag{11}$$

Take some $x_2' \,\epsilon\, X_2$ distinct from x_2. For a given domain value $x_1' \,\epsilon\, X_1$, either $(x_1', x_2') \,\epsilon\, CB$ or not. If there is no such exemplar, then equation (11) is satisfied directly. Otherwise $(x_1', x_2') \,\epsilon\, CB$ and it must be shown there is some other exemplar (x_1'', x_2'') in the case-base such that $\sigma(x, x_1') < \sigma(x, x_1'')$, which would prevent (x_1', x_2') being a nearest neighbour. Now from equation (10) it follows firstly:

$$\exists x_1 \,\epsilon\, X_1 \cdot (x_1, x_2) \,\epsilon\, CB \tag{12}$$

and additionally for some such value of x_1 and for any other exemplar $(x_1'', x_2'') \,\epsilon\, CB$ then $\sigma(x, x_1) \geq \sigma(x, x_1'')$. Therefore specifically:

$$\sigma(x, x_1) \geq \sigma(x, x_1') \tag{13}$$

But from equation (9), then $f(x_1) \neq f(x_1') \to \sigma(x, x_1) \neq \sigma(x, x_1')$. Since in this case $x_2 \neq x_2'$ then clearly $f(x_1) \neq f(x_1')$ by equation (1). Hence in addition to equation (13) we have $\sigma(x, x_1) \neq \sigma(x, x_1')$ and:

$$\sigma(x, x_1') < \sigma(x, x_1) \tag{14}$$

Hence by equations (12) & (14) (x_1, x_2) is an exemplar satisfying equation (11), and (11) is satisfied under all circumstances. **b) Only If.** Assume that equation (9) does not hold and therefore there is some $f' \,\epsilon\, F$ and some x, x' & x'' s.t.:

$$f(x') \neq f(x'') \wedge \sigma(x, x') = \sigma(x, x'') \tag{15}$$

Therefore consider the case-base $CB = \{(x', f(x')), (x'', f(x''))\}$. Since $f(x') \neq f(x'')$ we have two distinct reports satisfying equation (10) with respect to x. Hence $|\{x_2 | (x_1, x_2) \,\epsilon\, NN(x, CB, \sigma)\}| > 1$, and equation (9) necessarily follows from the case where $\sqsupseteq_I$ is admissible with respect to σ and F. □

3 Consistency of case-based learning algorithms

The similarity measure σ has been defined above only as a binary function from the space $(X_1 \times X_1 \to [0, 1])$; the property of *consistency* motivates some minimum constraints on a useful similarity measure. A consistent learner is one which, having seen some training sample, is able to correctly reproduce the examples it has seen in that sample. Although it is not optimal in domains where noise is expected, consistency is clearly a desirable property where it is assumed, as here, that the exemplars available to the system are error free. Since we all assume that the functions being represented are defined on a finite domain, consistency is sufficient to guarantee that as more training examples are seen the system will eventually converge to a good approximation of the target function [2, Chs 3 & 4]. This holds even in the worst case where the system is able to make little or no suitable generalisation of the seen examples. Since consistency is a property of learning *algorithms*, we must state explicitly the case-based learning algorithm we wish to consider. A family of the simplest such algorithms is defined below:

Definition 7 $CB1(\sigma)$ **Learning Algorithm for Case-Based Classifiers**

```
set CB = ∅
for i = 1 to m do
     set CB = CB ∪ {(x_i, f(x_i))}
set CB1(σ)(s̄) = h_⟨CB,σ⟩
```

These simple algorithms learn by adding each and every member of the training sample $\overline{s}$ (a series of m examples of the target function $(x_i, f(x_i))$) to the case-base and 'hypothesise' an approximation to the target function defined in terms of the case-base and a single fixed similarity measure σ. One of the simplest ways of ensuring $CB1(\sigma)$ is a consistent learning algorithm is to constrain σ so that the system will always retrieve an examplar from the training sample whenever that exemplar is presented as a query. The following definition is related to this intuition and is necessary and sufficient to ensure the consistency of $CB1(\sigma)$, a claim also proved below. The proof of Theorem 9 assumes that the case-base semantics of the previous section have been instantiated with an admissible preference relation.

Definition 8 Predictivity of a Similarity Measure with respect to a function space F **and preference ordering** $\sqsupseteq$**.** *A similarity measure is predictive of a function space* $F \subseteq (X_1 \to X_2)$*, iff for a function* $f \in F$*:*

1. *For any pair of problem instances* $x, x' \in X_1$*, the similarity of* x *to* x' *will exceed the similarity of* x *to itself only where* $f(x) = f(x')$*:*

$$\forall f \in F \cdot \forall x, x' \in X_1 \cdot \sigma(x, x') > \sigma(x, x) \to f(x) = f(x') \tag{16}$$

2. *For any pair of problem instances* $x, x' \in X_1$*, the similarity of* x *to* x' *will equal that of* x *to itself only if* $f(x')$ *is either equal to* $f(x)$ *or it does not precede* $f(x)$ *in the preference relation.*

$$\forall f \in F \cdot \forall x, x' \in X_1 \cdot \sigma(x, x') = \sigma(x, x) \to f(x) = f(x') \vee f(x') \not\sqsupseteq f(x)) \tag{17}$$

Theorem 9 Consistency of $CB1(\sigma)$**.** *For a space of functions* $F \subseteq (X_1 \to X_2)$*,* $CB1(\sigma)$ *is a consistent learning algorithm for* F *if and only if the chosen similarity measure* σ *is predictive of* F*.*

Proof: **a) Sufficiency.** Assume that a similarity measure σ is predictive of F, satisfying equations (16) & (17). For some example $(x, f(x))$ in the training sample, denote the members of $NN(x, CB, \sigma)$ by $(x_i^{NN}, f(x_i^{NN}))$. The definition of $CB1(\sigma)$ indicates that $(x, f(x))$ will be a member of CB and hence $\sigma(x, x_i^{NN}) \geq \sigma(x, x)$. Consider first the case where the inequality is strict. Hence $f(x) = f(x_i^{NN})$ for any of the nearest neighbours by equation (16). Therefore, whichever exemplar is the upper universal bound by the preference relation $\sqsupseteq$, $h_{\langle CB,\sigma\rangle}(x) = f(x)$. Hence in this case $CB1(\sigma)$ is a consistent learning algorithm. Assume instead that $\sigma(x, x_i^{NN}) = \sigma(x, x)$. By equation (17), either $f(x) = f(x_i^{NN})$ or $f(x_i^{NN}) \not\sqsupseteq f(x)$. For any nearest neighbours s.t $f(x_i^{NN}) \not\sqsupseteq f(x)$, clearly $f(x_i^{NN})$ cannot be the upper universal bound required by Definition 4. Therefore there must be at least one preferred nearest neighbour s.t. $f(x_i^{NN}) \sqsupseteq f(x)$. For any such exemplar necessarily $f(x_i^{NN}) = f(x)$, including the case where $(x_i^{NN}, f(x_i^{NN})) = (x, f(x))$. Hence for any element of the training sample, we have $h_{\langle CB,\sigma\rangle}(x) = f(x)$ and hence the algorithm is consistent. **b) Necessity.** Assume that equation (16) does not hold. Therefore there is some function $f \in F$ for which there are a pair of domain values x and x' s.t:

$$\sigma(x, x') > \sigma(x, x) \wedge f(x) \neq f(x')$$

Consider the training sample $\langle(x, f(x)), (x', f(x'))\rangle$. Clearly $CB1(\sigma)$ produces a hypothesis $\langle CB, \sigma\rangle$ s.t. $h_{\langle CB,\sigma\rangle}(x) = f(x')$, which is an inconsistent hypothesis. Assume alternatively equation (17) does not hold. Hence there are some f, x, x' s.t.

$$\sigma(x, x') = \sigma(x, x) \wedge f(x) \neq f(x') \wedge f(x') \sqsupseteq f(x)$$

Again, consider the hypothesis $\langle CB, \sigma\rangle$ produced by $CB1(\sigma)$ on the training sample $\langle(x, f(x)), (x', f(x'))\rangle$. Since x is equally similar to itself and to x', and $f(x')$ is preferred to $f(x)$ in the preference ordering, then again $h_{\langle CB,\sigma\rangle}(x) = f(x')$; again this is an inconsistent hypothesis. □

Example 10 Simple Concept Learning.

As above, we illustrate this result with reference to the special case of classification systems. In [4] we prove a version of Theorem 9 in terms of the following definition of 'special' predictivity for classification functions:

$$\forall f \,\epsilon\, F \cdot \forall x, x' \,\epsilon\, X_1 \cdot \sigma(x, x') \geq \sigma(x, x) \rightarrow f(x) = 1 \rightarrow f(x') = 1 \qquad (18)$$

$$\forall f \,\epsilon\, F \cdot \forall x, x' \,\epsilon\, X_1 \cdot \sigma(x, x') > \sigma(x, x) \rightarrow f(x) = 0 \rightarrow f(x') = 0 \qquad (19)$$

The following result re-expresses the special case [4, Thm 5] as a corollary of the general framework presented here:

Corollary 11 Consistency of Case-based classifiers *$CB1(\sigma)$ is a consistent learning algorithm for space of classification functions $F \subseteq (X_1 \rightarrow \{0, 1\})$ iff the similarity measure σ is predictive of F according to the 'special' definition of equations (18) & (19).*

Proof: Taking $X_2 = \{0, 1\}$ and $\sqsupseteq = \{(0, 0), (0, 1), (1, 1)\}$, then equations (16) & (17) become

$$\forall f \,\epsilon\, F \cdot x, x' \,\epsilon\, D_N \cdot \sigma(x, x') > \sigma(x, x) \rightarrow f(x) = f(x') \qquad (20)$$

$$\forall f \,\epsilon\, F \cdot x, x' \,\epsilon\, D_N \cdot \sigma(x, x') = \sigma(x, x) \rightarrow [f(x) = f(x') \vee (f(x') = 1 \wedge f(x) = 0)] \qquad (21)$$

By Theorem 9, $CB1(\sigma)$ will be a consistent learning algorithm for F iff equations (20) & (21) are satisfied. Hence it must be shown that a similarity measure σ satisfies (20) & (21) iff equations (18) & (19) are satisfied. **a) Only if.** Assume f, x & x' s.t. $\sigma(x, x') \geq \sigma(x, x)$ and $f(x) = 1$. Either the inequality is strict or the similarities are equal. In the case of a strict inequality, then $f(x') = 1$ by equation (20). Where the similarities are equal, and also $f(x) = 1$, then clearly equation (21) allows only $f(x') = 1$. Hence (18). Assume f, x & x' s.t. $\sigma(x, x') > \sigma(x, x)$ and $f(x) = 0$. From equation (20), immediately $f(x') = 0$. Hence (19). **b) If.** Assume x & x' s.t. $\sigma(x, x') > \sigma(x, x)$. For a given $f \,\epsilon\, F$, either $f(x) = 0$ or $f(x) = 1$. If $f(x) = 0$ then from (19) we have $f(x') = 0$ also. If $f(x) = 1$ then also $f(x') = 1$ by equation (18). Hence (20). Finally assume x & x' s.t. $\sigma(x, x') = \sigma(x, x)$ and some $f \,\epsilon\, F$ s.t. $f(x) = 1$ and $f(x') = 0$. But from (18), $f(x') = 1$ giving a contradiction. Hence (21). □

4 Conclusions

This paper has attempted to answer the question of how the knowledge content of a case memory system might be formalised. We have presented a view in which the case memory system is interpreted according to the semantics given as a function which approximates to some ideal mapping between input and output. The results presented explore a generalisation of the generally accepted decision function for a case-based classifier to a more general class of case-based systems which generate output from an arbitrary set of output values. The paper gives necessary and sufficient conditions for the well-definedness of our semantics and also for the consistency of case-based learning algorithms within this framework.

The functional view of knowledge content has a number of benefits, notably that it allows the error in the case memory's knowledge to be quantified straightforwardly in a way that is compatible with the assumptions of computational learning theory, allowing us to appeal to more general results in machine learning. We are finding that this allows some progress in understanding analytically the learning behaviour of various case-based reasoning systems [4]. The formalisation presented here might also be of use in emphasising the insight of Wess & Globig [9] that the knowledge content of a case memory system rests in the similarity measure as well as the stored cases.

Future work will attempt to make use of this framework in a model of case-based reasoning systems suitable for simple instances of the design task. We hope to make progress in understanding the nature and sources of error in the operation of a case memory system, and to develop these insights in a way that allows claims about the performance of case-based reasoning systems to be rigourously stated and proven.

Acknowledgements The first author is funded by an EPSRC grant and receives additional support under the CASE award scheme from Logica Cambridge Ltd.

References

[1] D W Aha, D Kibler, and M K Albert. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.

[2] M Anthony and N Biggs. *Computational Learning Theory.* Cambridge University Press, 1992.

[3] A M Dearden and M D Harrison. The engineering of case memory systems. *submitted to the Journal of Intelligent Information Systems.*

[4] A D Griffiths and D G Bridge. On concept space and hypothesis space in case-based learning algorithms. Submitted to *ECML-95: Eighth European Conference on Machine Learning 1995.* Available by WWW via `http://dcs2.cs.york.ac.uk:9876/isg/tony/tonydox.html`.

[5] K P Jantke. Case-based learning in inductive inference. In *COLT92 Proceedings of 5th ACM Worshop on Computational Learning Theory, July 92*, pages 218–223. ACM Press, 1992.

[6] K P Jantke and S Lange. Case-based representation and learning of pattern languages. In *EWCBR-93 Working Notes of the first European Workshop on Case-Based Reasoning*, volume 1, pages 139–144. University of Kaiserslautern, 1993.

[7] P Koton and M P Chase. Knowledge representation in a case-based reasoning system: Defaults and exceptions. In R J Brachman and H J Levesque, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, pages 203–211. Morgan Kaufmann, 1989.

[8] M Mehl. Retrieval in case-based reasoning using preferred subtheories. In G Brewka, K P Jantke, and P Schmitt, editors, *Nonmonotonic and inductive logic (2nd International Workshop)*, pages 284–297. Springer Verlag, 1993.

[9] S Wess and C Globig. Case-based and symbolic classification algorithms - A case study using version space. In S Wess, K-D Althoff, and M M Richter, editors, *Topics in CBR: EWCBR-93 - First European Workshop on Case-Based Reasoning, Kaiserslautern, Germany, November '93*, LNCS 837, pages 77–91. Springer-Verlag, 1994.