ELSEVIER Theoretical Computer Science 163 (1996) 161-176

Theoretical Computer Science

A framework for a theory of automated learning

E. Hausen-Tropper

Université du Québec à Montréal, Département d'informatique, PO Box 8888 Station Centre-Ville, Montreal, Que., Canada H3C 3P8

Received December 1993; revised July 1995
Communicated by M. Nivat

Abstract

This paper describes a topological theory of learning which characterizes learning as a homotopy (which is a function of the deformation of a space in time) from an original knowledge base to one augmented by learned knowledge. The knowledge representation, the algorithms and the theory are based on the cognitive theories of Jean Piaget, the psychological theories of Hebb, as well as the ethological theories of Merzenich and Kaas. The choice of topology as a descriptive and predictive theory was inspired by the neurological aspects of synaptic routing under learning, which preserves the original continuity of the routes. Topology was chosen as a basis for a theory of learning as it is the study of space invariants which preserve the structure and continuity of a space under "stretchings and deformations". Topology also offers a theoretical description of learning which does not revert to second-order logic. It has been shown that learning can be characterized by second-order logic, but any algorithm based on it is NP-complete.

1. Introduction

Machine learning has no fundamental theory outside of Michalski's Conceptual Clustering [10, 11] which is based on topological simplexes. Conceptual Clustering is an algorithm of classification based on concepts rather than on some quantitative or qualitative distance measure, as in mathematical clustering theories. The clustering is obtained through maximally disjunctive coverings of conjunctive normal forms. These coverings are simplexes which can be described by topologies on the space of concepts. Michalski's theory is applied to concepts, which he describes as objects in some context.

In this paper, a theory of learning, inspired by my knowledge representation, VMS, and learning algorithm, GANDALF [2], will be expounded. It is based on the cognitive theory of Piaget [12-16], the psychological theory of Hebb [5, 6], the ethological theory of Merzenich and Kaas [9] and the neurological theory of Schmitt et al. [17], and Hawkins and Kandel [4]. In the algorithm GANDALF, learning is

SSDI 0304-3975(95)00143-3

a process that takes place because of different triggers, depending on the class to which the concept belongs, the context and the motivation for learning. Establishing a theory of learning has the merit of determining when an augmentation of knowledge, through acquisition, can actually be described as learning.

This theory is capable of modelling learning and will eventually be capable of predicting degrees of learning through a metric on the topological spaces. This could lead to a totally new approach to the building of self-modifying expert systems, by predicting how far and how long the system could be viable. Eventually, the research will lead to a definition and characterization of machine creativity and intuition.

The research by neurologists on human learning has shown that the synapses formed during learning respect the polymorphy of the original state of the neuronic routes. That is, learning is merely a stretching or deformation of the pattern of the original circuitry.

Topology as a theory of learning was prompted by the mentioned neurological research, as well as my learning algorithms which could be described by topological invariants under unions and intersections. A topology is a space made up of all unions and intersections of its sets, as well as the null set and itself. It is the study of space invariants which preserve the structure and continuity of a space under "stretchings and deformations".

Thus, machine learning can be characterized by topological invariants, in fact, learning can be seen as a homotopy (deformation preserving the structure and continuity of the space) on an existing knowledge base.

In a previous paper [3] a knowledge representation for learning called a VMS, for variable multilinked schema, was defined and a learning system called GANDALF was implemented. In this paper, it will be shown that learning can be characterized by a continuous time deformation of an original knowledge base to an enriched knowledge base incorporating the learned elements. Learning will, thus, be shown to be a homotopy from some basic knowledge to enriched knowledge through acquisition.

1.1. Basics of topology applicable to learning

The following is taken from Dugundji [1]. It represents the bare essentials of Topology applicable to a theory of learning.

Definition 1.1. Given a set X, a topology $\mathscr{T}$ in X is a family of subsets of X such that:
(1) each union of members of $\mathscr{T}$ is in $\mathscr{T}$,
(2) each finite intersection of members of $\mathscr{T}$ is in $\mathscr{T}$,
(3) $\emptyset$ and X are in $\mathscr{T}$.

Definition 1.2. An open set in topology $\mathscr{T}$ is a member of $\mathscr{T}$.

Definition 1.3. A closed set in topology $\mathscr{T}$ is a set which is not open.

Theorem 1.1. *Given any family $S = \{A_a \mid A_a \text{ in } X\}$, there exists a unique, smallest topology $\mathscr{T}(S)$ containing S:*

$$\mathscr{T}(S) = \{\emptyset, X, \textit{all finite intersections of } A_a, \textit{all unions of finite intersections of } A_a\}.$$

Theorem 1.2. *A map $f: X \to Y$ is continuous iff $f^{-1}: \mathscr{T}_y \to \mathscr{T}_x$.*

Note: This holds if every open (closed) set in $\mathscr{T}_y$ maps into an open (closed) set in $\mathscr{T}_x$.

Theorem 1.3. *A map, $f: X \to Y$ is continuous iff f^{-1} $(U \mid U$ subbasis of $Y)$ is open in X.*

Definition 1.4. For any Cartesian product, $\prod_a Y_a$, of sets, $\{Y_a \mid a \in A\}$, the projection, p_b, along basis b is defined by

$$p_b; \prod_a Y_a \to Y_b \quad \text{for } b \in A.$$

Definition 1.5. For any family $\{(Y_a, \mathscr{T}_a) \mid a \in A\}$, $\prod_a Y_a$ is a topological space having for subbasis all

$$\langle U_b \rangle = p^{-1} b(U_b)$$

such that

$$p_b^{-1}(U_b): U_b \to P_a \mathscr{T}_a \quad \forall\, U_b \in \mathscr{T}_b,\ \forall\, b \in A$$

with the Cartesisn product topology $\prod_a \mathscr{T}_a$.

Definition 1.6. A map of two variables $f(x, y) = z$ can be regarded as a map on a on Cartesian product,

$$f: X \times Y \to Z,$$

or as a family of maps,

$$f_x: Y \to Z$$

with X as parameter.

Definition 1.7. A free union $X + Y$ of topological spaces X and Y is a topological space $(X \cup Y)$, where a set B in $X + Y$ is closed (or open) iff $B \cap X$ and $B \cap Y$ are closed (or open).

Definition 1.8. Given a family $\{A_a \mid A_a \text{ in } X\}$, such that

(1) for all a, b, the topologies of A_a, A_b agree on $(A_a \cap A_b)$,

(2) each open (closed) set in $(A_a \cap A_b)$ is open (closed) in A_a, A_b then the weak topology, is

$$\mathscr{T}_a = \{U \text{ in } X \mid \forall a\colon U \cap A_a \text{ is open (closed) in } A_a\}.$$

Definition 1.9. Given $f, g: X \to Y$, where X and Y are topological spaces, f is homotopic to g (written $f \cong g$) if there exists a continuous

$$F: X \times I \to Y \quad \text{for } I = \{t \mid 0 \leqslant t \leqslant 1\}$$

such that

$$F(x,0) = f(x),\ F(x,1) = g(x) \quad \forall x \in X.$$

A homotopy is, then, a time deformation of a map into another map.

1.2. Definition of variable multilinked schemata or VMS

The following is a review of the knowledge representation VMS.

A VMS is a multidimensional data structure consisting of schemata with slots whose possible values are elements of a set. Each element of these sets can be a conjunction of elements. The first four slots represent triggers, class, context and motivation. The fifth slot is a pre-set action for a particular linked 4-tuple of slot values. The sixth slot is a learned action. In the original knowledge base, only the first five slots can be filled. The sixth slot will be filled by learning. It is to be noted that the unnecessity of new learning, or simply following the pre-set action in the fifth slot, can be viewed as a null learning set. This is important in the concept of learning as a partially continuous function.

We note that either the fifth slot is filled and the sixth slot is empty in a linked list, or the sixth slot is filled and the fifth slot is empty.

Definition 1.10. A variable multilinked schema is defined as follows:

$$\begin{aligned}
&F_i \\
&\quad s_{1i}\colon\ \{s_{1i}^1, \ldots, s_{1i}^n\} \\
&\quad s_{2i}\colon\ \{s_{2i}^1, \ldots, s_{2i}^m\} \\
&\qquad \ldots \\
&\quad s_{6i}\colon\ \{s_{6i}^1, \ldots, s_{6i}^r\}
\end{aligned}$$

where

s_{ji} is the name of the slot (not used in this paper) and j: $1 \leqslant j \leqslant 6$

s_{ji}^k is the kth value of a set of possible values for slot j, schema i.

Then we have:

1st subscript refers to the slot number,

2nd subscript refers to the schema number,

1st superscript refers to the position of that value in the set of possible values for that slot.

Any of these slot values, s_{ji}^k, may be a set representing the conjunction of values.

The sets above are formed in a horizontal fashion; but each value for each slot is connected to one or more values of all the other slots. It is therefore a multilinked list.

1.3. *Learning as a continuous function in a discrete event system*

Although Gandalf is a reasoning and learning system triggered by discrete input events, it simulates a machine capable of continuous action. It is to be noted that the knowledge base of the system simulates the brain of a continuously operating animal (whether thinking, moving, standing still, dreaming or breathing).

As mentioned before, the Gandalf system runs continuously either triggered into a pre-set action (specified in slot five), or triggered into a new learning state (by filling slot six), or finding possible new links from combinations of similar slot values, or destroying learned actions not used for a long time. As a modification of its knowledge base, all these actions can be viewed as learning.

At any time t, Gandalf can be in any of the following states:

(1) following the pre-set action in slot five triggered by an input,

(2) adding a new learned action in slot six triggered by an input,

(3) removing a learned slot value in slot six, or removing a whole schema upon reaching a specific time stamp in the sixth slot,

(4) randomly selecting schemata or slots with similar values and combining them into new learned links or a new schema.

The transitions between any of these four states are the discontinuities in the partially continuous learning function.

2. A topological theory of learning

In the following, it will be shown that learning can be characterized as a continuous-time deformation, or homotopy, from a basic knowledge base to an enriched knowledge base containing the acquired concepts.

2.1. *VMS's as topological spaces*

For the following topological treatment, we use intersections and unions in lieu of conjunctions and disjunctions, since we treat our elements as sets. Semantically, the meaning will remain that of conjunction and disjunction.

Definition 2.1. A general VMS (without its slot labels) is a doubly linked data structure which is a set of the form:

$$\begin{aligned}F_i = ((&s_{1i}^1, \ldots, s_{1i}^n),\\ (&s_{2i}^1, \ldots, s_{2i}^m),\\ &\ldots\\ (&s_{6i}^1, \ldots, s_{6i}^g))\end{aligned}$$

of disjunctive elements, s_{ji}^k, where each s_{ji}^k can be a set of conjunctive elements.

Theorem 2.1. *For a fixed schema F_i, let $\{S_{ij}\}$ be a family of subsets of F_i such that:*

$S_{ij} = \{s_{1i}^k, s_{2i}^1, \ldots, s_{6i}^q\}$ *is true.*

A set which is a linked list in the knowledge base is considered true. Then $\{S_{ij}\}$ is a topological space in F_i.

Proof.

(1) If $\{s_{1i}^k, s_{2i}^l, \ldots, s_{6i}^q\}$ is true, then from set-theoretic considerations any union of members of $[S_{ij}]$ is true:

$S_{it} \cup S_{iu} \in \{S_{ij}\} \quad t, u \in j.$

(2) If $\{s_{1i}^k, s_{2i}^l, \ldots, s_{6i}^q\}$ is true, then from set-theoretic considerations any finite intersection of members of $\{S_{ij}\}$ is true:

$S_{it} \cap S_{iu} \in \{S_{ij}\} \quad t, u \in j.$

(3) $(\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset) \in F_i$ is the no action, no learning state for all i. Therefore: $(\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset) \in \{S_{ij}\}$ for some j. Also, $F_i = \bigcup_j S_{ij}$ over all j is true from (1) above. Therefore, $F_i \in \{S_{ij}\}$. Therefore, from Definition 1.1, $\{S_{ij}\}$ is a topology in F_i. □

Note: The no-action state is not a no-thinking stage. In that time garbage collection can be executed, or new links can be added by randomly picking combinations of existing schemata with similar slot values.

Theorem 2.2. *Each linked list $\{s_{1i}^k, s_{2i}^l, \ldots, s_{6i}^q\}$ in the knowledge base is an open set of F_i. Each disconnected list in $\{S_{ij}\}$ is a closed set of F_i.*

Proof. From Definitions 1.2 and 1.3, since each linked list is a member of S_{ij} and each disconnected list is not, the proof follows. □

Theorem 2.3. *For any family of VMS, $\{S_{ij} \mid S_{ij} \in F_i\}$, in the knowledge base, $\prod_i S_{ij}$, is a topological space.*

Proof.

(1) By Theorem 1.3, for each $S_{ij} \in F_i$ there exists a topology $\mathcal{T}_i$ such that $\{S_{ij}, \mathcal{T}_i\}$ is a topological space.

(2) By Definition 1.4, for any family of topological spaces $\{S_{ij}, \mathcal{T}_i\}$, $\{\prod_i S_{ij}, \prod_i \mathcal{T}_i\}$ is a topological space.

(3) Furthermore, by Definition 1.8, if

$\langle O_k \rangle = p_k^{-1}(O_k)$

such that

$$p_k^{-1}(O_k): O_k \to \prod_i \mathcal{T}_i \quad \forall\, O_k \in \mathcal{T}_i$$

then $\{\langle O_k \rangle \mid \langle O_k \rangle \in \prod_i \mathcal{T}_i\}$, forms a subbasis for the Cartesian product topology, $\prod_i \mathcal{T}_i$. $\square$

This implies that members of the Cartesian product topology are composed of projections of the product topology into members of the original topologies.

This theorem states that the open sets, or linked lists, of the conjunction of schemata is composed of linked lists of the original schemata.

2.2. *Intra-schema learning as a homotopy from one base of knowledge to another*

In intra-schema learning, if the learned actions are results of conjunctions and disjunctions of linked lists within one schema, we have the following.

Lemma 2.1. *Let S_n be the set of all possible values of the nth slot of the schema, F_i. Then*

$$\{(s_{1i}^j, s_{2i}^k, s_{3i}^l, s_{4i}^m, s_{5i}^o, s_{6i}^p)\} \in (S1 \times S2 \times S3 \times S4 \times S5 \times S6) = \mathbf{V}$$

$$\{(s_{1i}^j, s_{2i}^k, s_{3i}^l, s_{4i}^m)\} \in (S1 \times S2 \times S3 \times S4) = \mathbf{S}$$

for a fixed i are topological spaces.

Proof. Proof follows directly from Theorem 1.1. $\square$

Lemma 2.2. *Let $\boldsymbol{J} = \{(s_1, s_2, s_3, s_4)\}$ be the set of nonnoisy values of the input into the system which triggers some schema(ta). Then $\boldsymbol{J}$ is a topological space.*

Proof. If the input is nonnoisy it is assumed to be correct. The proof follows that of Theorem 1.3. The open sets are input sets which match linked lists in the knowledge base. $\square$

Theorem 2.4. *Let $\boldsymbol{J} = \{(s_1, s_2, s_3, s_4)\}$ be the set of nonnoisy values of the input. Let $\boldsymbol{S} = (S1 \times S2 \times S3 \times S4)$ be the set of 4-tuples representing the first four slots of schema F_i then*

$$\boldsymbol{X} = \boldsymbol{J} + (S1 \times S2 \times S3 \times S4) \quad \text{for a fixed } i$$

is a free union with weak topology:

$$\mathcal{T}_i = \{\textit{open (closed) } U \textit{ in } \boldsymbol{X} \mid U \cap \boldsymbol{J} \textit{ is open (closed) in } \boldsymbol{J}, \\ \textit{and } U \cap \boldsymbol{S} \textit{ is open (closed) in } \boldsymbol{S}.\}$$

Proof. Let U be a 4-tuple in $(J \cup S)$.

(1) $U \cap \boldsymbol{S}$ is open in S if it represents a linked list in some schema. $U \cap \boldsymbol{J}$ is open in $\boldsymbol{J}$ if it is nonnoisy and matched linked values of that schema. In that case U will be a linked 4-tuple in $(J \cup S)$ and, therefore, an open set of $(J \cup S)$.

(2) The same argument holds for closed sets.

(3) If U is open in $(J \cup S)$ it is a linked 4-tuple and therefore will be linked in $\boldsymbol{S}$ and in $\boldsymbol{J}$.

Therefore, from Definitions 1.6 and 1.7, $(\boldsymbol{J} + \boldsymbol{S})$ is a free union.

We now show that intra-schema learning is a continuous-time deformation from a pre-set action in a fixed schema, to a new action in that same schema.

Lemma 2.3. *$f: \boldsymbol{J} + (S1 \times S2 \times S3 \times S4) \to (S1 \times S2 \times S3 \times S4 \times S5 \times S6)$. That is $f: (\boldsymbol{J} + \boldsymbol{S}) \to V$ such that*

$$f(x) = (s_{1i}^k, \ldots, s_{4i}^n, s_{5i}^o, \emptyset) \quad \textit{for a linked list in } F_i$$

$$= (s_{1i}^k, \ldots, s_{4i}^n, \emptyset, \emptyset) \quad \textit{for a disconnected list in } F_i.$$

$$g: \boldsymbol{J} + (S1 \times S2 \times S3 \times S4) \to (S1 \times S2 \times S3 \times S4 \times S5 \times S6)$$

such that

$$g(x) = (s_{1i}^k, \ldots, s_{4i}^n, s_{5i}^o, s_{5i}^o) \quad \textit{for a linked list in } F_i$$

$$= (s_{1i}^k, \ldots, s_{4i}^n, \emptyset, s_{6i}^p) \quad \textit{for a disconnected list in } F_i.$$

Then f, g are functions from topologicakl space $(\boldsymbol{J} + \boldsymbol{S})$ into topological space V.

Proof. (1) For any linked list $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n) \in \boldsymbol{S}$ there is a s_{5i}^o which is connected to it and represents the pre-set action. In this case the learned action in S_6 will be $\emptyset$.

(2) Since $(\boldsymbol{J} + \boldsymbol{S})$ is a free union, $(s_1, s_2, s_3, s_4) \cap (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n)$ is a linked list in $(\boldsymbol{J} + \boldsymbol{S})$ if (s_1, s_2, s_3, s_4) is a linked list in $\boldsymbol{J}$ and $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n)$ is a linked list in $\boldsymbol{S}$. In this case

$$f(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, \emptyset) \quad \text{for some unique } s_{5i}^o$$

$$g(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, s_{5i}^o) \quad \text{for } x \in (s_1, s_2, s_3, s_4) \cap (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n).$$

(3) For a disconnected list (s_1, s_2, s_3, s_4) if the triggered lists in the schema are combined through conjunctions and disjunctions, they give a specific s_{6i}^p such that

$$g(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, \emptyset, s_{6i}^p)$$

$$f(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, \emptyset, \emptyset)$$

for $x \in (s_1, s_2, s_3, s_4) \cap ((s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n), \ldots, (s_{1i}^q, s_{2i}^r, s_{3i}^s, s_{4i}^t))$. $\square$

Theorem 2.5. *Let*

$$f: \boldsymbol{J} + \boldsymbol{S} \to V$$

$$g: \boldsymbol{J} + \boldsymbol{S} \to V \quad \textit{S and V as in Lemma 2.3.}$$

Then there exists a continuous

$$F: (\boldsymbol{J} + \boldsymbol{S}) \times I \to V$$

where $I = [0, 1]$ *such that*

$$F(x, 0) = f(x) \text{ and } F(x, 1) = g(x) \quad \forall x \in (\boldsymbol{J} + \boldsymbol{S})$$

Therefore f is homotopic to g.

Proof. Given any member U of the topology for space V, it will have one of the following forms:

(1) $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, \emptyset, \emptyset)$,
(2) $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, \emptyset, s_{6i}^p)$,
(3) $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, \emptyset)$,
(4) $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, s_{5i}^o)$.

If it is of form (1) then $F^{-1}(U) \to (x, 0)$ for $x \in (\boldsymbol{J} + \boldsymbol{S})$, x a disconnected list. If it is of form (2) then $F^{-1}(U) \to (x, 1)$ for $x \in (\boldsymbol{J} + \boldsymbol{S})$, x a disconnected list, where

$$x \in (s_1, s_2, s_3, s_4) \cap ((s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n), \ldots, (s_{1i}^q, s_{2i}^r, s_{3i}^s, s_{4i}^t)).$$

If it is of form (3) then $F^{-1}(U) \to (x, 0)$ for $x \in (\boldsymbol{J} + \boldsymbol{S})$, x a linked list. If it is of form (4) then $F^{-1}(U) \to (x, 1)$ for $x \in (\boldsymbol{J} + \boldsymbol{S})$, x a linked list, where $x \in (s_1, s_2, s_3, s_4) \cap (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n)$. That is, any open set in V maps into an open set in $((\boldsymbol{J} + \boldsymbol{S}) \times I)$. Therefore, by Theorem 2.2, f is homotopic to g. □

2.3. Inter-schemata learning as a homotopy from one base of knowledge to another

In inter-schemata learning, if the learned actions are the results of conjunctions of linked lists in more than one schema we proceed as follows. In Theorem 2.3, we showed that the resulting schemata obtained through these conjunctions form a topological space.

Lemma 2.4. *Let R_j be the set of possible values for the jth slot of a set of schemata $\{F_t\}$ in the knowledge base. For example:*

$$\{s_{jq}^k, s_{jr}^l, s_{js}^m\} \in R_j \quad \text{for } s_{jq}^k \in F_q, \; s_{jr}^l \in F_r, \; s_{js}^m \in F_s.$$

Let $T = (R_1 \times R_2 \times R_3 \times R_4) = \prod_1^4 R_j$. *Let* $W = (R_1 \times R_2 \times R_3 \times R_4 \times R_5 \times R_6) = \prod_1^6 R_j$. *Then T and W are topological spaces.*

Proof. Proof follows from Definition 1.6 and Lemma 2.1. □

Lemma 2.5. *Let $\{f_i\}$ be a family and $f_i, g: \boldsymbol{J} + T \to W$ ranging on a set of schemata $\{F_t\}$ such that for any $x \in (\boldsymbol{J} + T)$*

$$f_i(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, \emptyset) \quad \textit{for a linked list in } F_i$$

$$= (\Lambda_t s_{1t}^k, \Lambda_t s_{2t}^l, \Lambda_t s_{3t}^m, \Lambda_t s_{4t}^n, \emptyset, \emptyset)$$

for a disconnected list in the KB spanning schemata $\{F_t\}$.

$$g(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, s_{5i}^o) \quad \textit{for a linked list in } F_i$$

$$= (\Lambda_i s_{1i}^k, \Lambda_i s_{2i}^l, \Lambda_i s_{3i}^m, \Lambda_i s_{4i}^n, \emptyset, \Lambda_i s_{5i}^o)$$

for a disconnected list in the KB spanning schemata $\{F_t\}$. Then $\{f_i\}$ and g are functions from space $(\boldsymbol{J} + \boldsymbol{T})$ to space W.

Proof. (1) The proof follows the same lines as in section (1) of Lemma 2.3.

(2) The proof follows the same line as in section (2) of Lemma 2.3.

(3) For a disconnected list (s_1, s_2, s_3, s_4) if the triggered lists in the schemata are combined by conjunctions to give:

$$g(x) = (\Lambda_i s_{1i}^k, \Lambda_i s_{2i}^l, \Lambda_i s_{3i}^m, \Lambda_i s_{4i}^n, \emptyset, \Lambda_i s_{5i}^o)$$

$$f_i(x) = (\Lambda_i s_{1t}^k, \Lambda_i s_{2t}^l, \Lambda_i s_{3t}^m, \Lambda_i s_{4t}^n, \emptyset, \emptyset)$$

for $x \in ((s_1, s_2, s_3, s_4) \cap (\Lambda_i s_{1i}^k, \Lambda_i s_{2i}^l, \Lambda_i s_{3i}^m, \Lambda_i s_{4i}^n))$. $\square$

We now show that inter-schemata learning is a continuous-time deformation from pre-set actions in a set of schemata, to a new action in a new schema.

Theorem 2.6. *Let $\{f_i\}$ be a family of maps $f_i: (\boldsymbol{J} + T) \to W$ such that for any $x \in (\boldsymbol{J} + T)$ spanning schemata $\{F_t\}$*

$$f_i(x) = ((s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, \emptyset) \textit{ for a linked list in } F_i$$

$$= (\Lambda_i s_{1t}^k, \Lambda_i s_{2t}^l, \Lambda_i s_{3ti}^m, \Lambda_i s_{4t}^n, \emptyset, \emptyset)$$

for a disconnected list in the KB spanning schemata $\{F_t\}$.

$$g(x) = (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, s_{5i}^o) \quad \textit{for a linked list in } F_i$$

$$= (\Lambda_i s_{1i}^k, \Lambda_i s_{2i}^l, \Lambda_i s_{3i}^m, \Lambda_i s_{4i}^n, \emptyset, \Lambda_i s_{5i}^o)$$

for a disconnected list in the KB spanning schemata $\{F_t\}$. Then there exists a family of continuous functions $F_i: (\boldsymbol{J} + T) \times I \to W$ where $I = [0, 1]$ such that

$$F_i(x, 0) = f_i(x) \quad \textit{and} \quad F_i(x, 1) = g(x) \quad \forall x \in (\boldsymbol{J} + T).$$

Therefore, f_i is homotopic to g for all i spanning set $\{F_t\}$ of schemata.

Proof. Given any member U of the topology for space W, it will have one of the following forms:

(1) $(\Lambda_i s_{1t}^k, \Lambda_i s_{2t}^l, \Lambda_i s_{3ti}^m, \Lambda_i s_{4t}^n, \emptyset, \emptyset)$,

(2) $(\Lambda_i s_{1i}^k, \Lambda_i s_{2i}^l, \Lambda_i s_{3i}^m, \Lambda_i s_{4i}^n, \emptyset, \Lambda_i s_{5i}^o)$,

(3) $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, \emptyset)$,

(4) $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n, s_{5i}^o, s_{5i}^o)$.

If it is of form (1) then $F_i^{-1}(U) \to (x, 0)$ for $x \in (\boldsymbol{J} + T)$, x a disconnected list. If it is of form (2) then $F_i^{-1}(U) \to (x, 1)$ for $x \in (\boldsymbol{J} + T)$, x a disconnected list, where $x \in (s_1, s_2, s_3, s_4) \cap ((s_{1o}^k, s_{2o}^l, s_{3o}^m, s_{4o}^n), \ldots, (s_{1j}^q, s_{2j}^r, s_{3j}^s, s_{4j}^t))$.

It it is of form (3) then $F_i^{-1}(U) \to (x, 0)$ for $x \in (\boldsymbol{J} + T)$, x a linked list. If it is of form (4) then $F_i^{-1}(U) \to (x, 1)$ for $x \in (\boldsymbol{J} + T)$, x a linked list, where $x \in (s_1, s_2, s_3, s_4) \cap (s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n)$ and $(s_{1i}^k, s_{2i}^l, s_{3i}^m, s_{4i}^n) \in F_i$. That is, any open set in W maps into an open set in $((\boldsymbol{J} + T) \times I)$. Therefore, by Theorem 1.2, f_i is homotopic to g for all i spanning $\{F_t\}$. $\square$

This theorem says that inter-schemata learning can be regarded as a time deformation from a set of pre-set actions to a learned pattern combining all these pre-set actions.

With Theorems 2.5 and 2.6 we have shown that learning can be characterized as a topological deformation from pre-set actions to learned actions.

3. Learning as a homotopy in the system GANDALF

In this section, we will first describe the learning algorithm, GANDALF, and then show that it follows the requirements of the above topological theory of learning. Having established that, we can affirm that learning in the system GANDALF can be described as a homotopy from a base of knowledge to a base augmented by the learning process.

3.1. The learning algorithm GANDALF

The input into the system is an ordered 4-tuple: $\boldsymbol{J} = (s_1, s_2, s_3, s_4)$. Any of these may be null or $\emptyset$. Any of these may be a set representing conjunctions of values for that slot: $s_k = (s_k^1, \ldots, s_k^m)$. From an inverse list of (slot values: schema numbers) we obtain a set of schemata with some matching values in the first four slots, if possible.

The reason for the partial match of slot values is the following. The process of recognition is not a simple matching algorithm between observed properties and stored properties of a concept. Very often the properties we use for recognition are only a small fraction of the total properties that are deemed necessary for positive identification (typical is the case of occlusion). We also identify very familiar animals, such as cats, by the pointed ears and whiskers only.

There are many possible combinations of matches between one or more of the input's slot values and the schemata's slot values, but fortunately most are not interesting for learning purposes. Three situations may arise, aside from finding no schema with "sufficient" matching values in any slot, giving the no-action, no-learning state $\Phi = (\Phi, \Phi, \Phi, \Phi, \Phi)$.

(1) Schemata are found with "sufficient matching values in the first (triggering) slot, and we obtain schemata, $F_1, \ldots, F_t$, whose first slot contains values that match some in the set s_1, and contain no values that are the negations of some values in the input, i.e., $s_{1i}^j \cap s_1 \neq 0$ for some j depending on i, for all i: $1 \leqslant i \leqslant t$,

$$s_{1i}^k \neq -s_{11} \quad \forall s_{1i}^k \in s_{1i},\ \forall s_{11} \in s_1,\ \forall i\colon 1 \leqslant i \leqslant t.$$

(2) No schema is found with sufficient matching values in the first (triggering) slot. In this case we pick schemata $F_1, \ldots, F_r$ which belong to the same class and treat the input as a new instance of these classes:

$$s_2 = s_{2i}^j \quad \text{for some } j \text{ depending on } i, \text{ for all } i\colon 1 \leqslant i \leqslant r,$$

$$s_2 \neq -s_{2i}^k \quad \text{for some } k \text{ depending on } i, \text{ for all } i\colon 1 \leqslant i \leqslant r.$$

(3) No schema is found with sufficient matching values in the first and second slots.

We show the possibilities for the first and second cases in the Table 1.

Through an elaborate mechanism based on Table 1, GANDALF triggers the appropriate schema or schemata for learning.

3.2. Learning algorithms in GANDALF

There are two types of learning, intra-schema and inter-schemata. The first occurs when only one schema is involved, while the second occurs when at least two

Table 1

Schemas match					
S1	*S2*	*S3*	*S4*	Schemas picked	Type of action or learning
✓	✓	✓	✓	F_1	Take preset action
✓	∅	∅	∅	$F_1, \ldots, F_n$	Identification problem Pick most salient features
✓	✓	∅	∅	Subset of $(F_1, \ldots, F_n)$	Identification problem Pick most salient features
✓	∅, ✓	× or	×	$(F_1, \ldots, F_n)$ or subset	Learn new context or motivation, form new links
✓	×	ψ	ψ	Match $S1 \to (F_1, \ldots, F_n)$ Match $S2 \to (F_i, \ldots, F_m)$	Learn new class of schema
×	✓	ψ	ψ	Match $S2 \to (F_i, \ldots, F_m)$	Learn new instance

Note: ✓ indicates a match between the input's slot values and the schemata's slot values. × indicates a nonmatch between the input's slot values and the picked schemata's slot values. ∅ indicates that no value is given for that slot in the input. ψ indicates the presence (matching or non-matching) of the given value for that slot.

schemata are involved. Within one schema, we consider first whether more than one value was matched for the first or second slots. If such is the case, we take the conjunction of all values in that slot, as well as of values in all other slots connected to it. This gives rise to the procedure INTRA-CONJ.

INTRA-CONJ: Input matching two or more values in first or second slot of one schema

If the input matches two or more values in the triggering, or class slot of one schema:

$$s_1 \cap s_{1i}^k \neq 0 \text{ and } s_1 \cap s_{1i}^l \neq 0$$

then we obtain the pre-set linked lists

$$(s_{1i}^k, \ldots, s_{5i}^x, \emptyset), (s_{1i}^l, \ldots, s_{5i}^y, \emptyset)$$

$\forall$ connections to s_{1i}^k and s_{1i}^l and take their conjunction

$$\Lambda x\, \Lambda y (s_{1i}^k, \ldots, s_{5i}^x, \emptyset) \wedge (s_{1i}^l, \ldots, s_{5i}^y, \emptyset)$$

That is, if s_{1i}^k is the first element of linked lists:

$$(s_{1i}^k, \ldots, s_{5i}^m, \emptyset), (s_{1i}^k, \ldots, s_{5i}^n, \emptyset), (s_{1i}^k, \ldots, s_{5i}^o, \emptyset)$$

and if s_{1i}^l is the first element of linked lists:

$$(s_{1i}^l, \ldots, s_{5i}^p, \emptyset), (s_{1i}^k, \ldots, s_{5i}^q, \emptyset)$$

then we obtain

$$(s_{1i}^k \wedge s_{1i}^l, \ldots, \emptyset, s_{5i}^m \wedge s_{5i}^n \wedge s_{5i}^o \wedge s_{5i}^p \wedge s_{5i}^q)$$

with $s_{5i}^? = \emptyset$ and the learned action:

$$(s_{5i}^m \wedge s_{5i}^n \wedge s_{5i}^o \wedge s_{5i}^p \wedge s_{5i}^q)$$

which is stored as

$$s_{6i}^? = ((k, l), \ldots, (m, n, o, p, q)).$$

In the case where we match only some of the slots, or when we match all first four slots but they are not linked in the schema, we take the disjunction of all values in one slot which are linked to the matched values. This gives rise to the procedure INTRA-DISJ.

INTRA-DISJ: *Input matching slot values disconnected within one schema*

When the input gives only the first or second slot, or when we have a match between the input and the schema on more than one slot, we must consider whether or not the values for these slots are connected in the schema. We recall that the values in each slot are connected to values in other slots of that schema. In the case of matching only one slot, or disconnected slot values, we have no one corresponding pre-set action.

If the values are connected within the schema, the action taken is the pre-set action.

If the values are for one slot only, or are not connected within the schema we take disjunctions of the involved linked lists.

For example, consider the following match between the input $(s_1, \ldots, s_4)$ and the schema F_i:

$$s_1 \cap s_{1i}^k \neq \emptyset, \quad s_2 \cap s_{2i}^l \neq \emptyset, \quad s_3 \cap s_{3i}^m \neq \emptyset, \quad s_4 \cap s_{4i}^n \neq \emptyset.$$

Furthermore, suppose that within F_i, the established connections are:

$$(s_{1i}^k, \ldots, s_{5i}^l, \emptyset), (s_{1i}^k, \ldots, s_{5i}^2, \emptyset), (s_{1i}^k, \ldots, s_{5i}^3, \emptyset)$$

$$(s_{1i}^o, s_{2i}^l, \ldots, s_{5i}^4, \emptyset)$$

$$(s_{1i}^p, \ldots, s_{3i}^m, \ldots, s_{5i}^5, \emptyset), (s_{1i}^q, \ldots, s_{3i}^m, \ldots, s_{5i}^6, \emptyset)$$

$$(s_{1i}^r, \ldots, s_{4i}^n, s_{5i}^7, \emptyset).$$

Some or all of the pre-set actions, s_{5i}^x may overlap. We learn the new action: $Vx\, s_{5i}^x$ where $x = 1, 2, 3, 4, 5, 6, 7$ in the above case. In general, we learn the action:

$$Vx\, s_{5i}^x \quad \text{where } x = \{a\}, \{b\}, \{c\}, \{d\}$$

and

$s_{5i}^{\{a\}}$ are connected to s_{1i}^k,

$s_{5i}^{\{b\}}$ are connected to s_{2i}^l,

$s_{5i}^{\{c\}}$ are connected to s_{3i}^m,

$s_{5i}^{\{d\}}$ are connected to s_{4i}^n.

Should any of $s_1, \ldots, s_4 = \emptyset$, we would omit the connection to the fifth slot for that entry. The learned pattern would be:

$$s_{6i}^? = (((k), (o), (p), (r)), \ldots, ((\{a\}), (\{b\}), (\{c\}), (\{d\}))).$$

In the case where more than one schema was triggered, we take the conjunction of the slot values of all the schemata matched. This gives rise to the procedure INTER-CONJ.

INTER-CONJ: Learning with more than one schema

Given a triggered set, $(F_r, \ldots, F_p)$, for learning, we apply inter-schemata learning to the learned patterns obtained from each schema using intra-schema learning. This is accomplished by taking the intersection across schemata for the learned action slot.

We learn:

$$(\Lambda_i s_{1i}^j, \Lambda_i s_{2i}^k, \Lambda_i s_{3i}^l, \Lambda_i s_{4i}^m, \emptyset, \Lambda_i s_{6i}^n) \quad \forall i \ni F_i \subseteq (F_r, \ldots, F_p)$$

where j, k, l, m, n depend on the particular i involved and with which we then form a new schema representing a new entity.

We can then affirm that learning in the system GANDALF is a homotopic process of going from an original knowledge base to one augmented by learning. As such, it is a time deformation of one base into another respecting the original polymorphy of the base.

4. Conclusion

In this paper we first briefly described the knowledge representation called variable multi-linked schema, VMS. This representation was a doubly linked data structure.

We then described how a topological theory of learning could be established by considering homotopies on a data structure of linked lists. In fact, a process of augmenting the data structure by adding certain conjunctions and disjunctions within one schema, or among more than one schema resulted in creating a deformation of the original structure which respected the requirements of a topological invariant. That invariant was a homotopy from the original structure to the new structure.

We then showed that the learning algorithm GANDALF operated through conjunctions and disjunctions, as described in the topological theory of learning. This established GANDALF as an algorithm which could be described by a topological theory of learning.

It was shown that learning can then be characterized as a homotopy from an original knowledge base to an enriched one through knowledge acquisition.

References

[1] J. Dugundji, *Topology*, Allyn and Bacon Series in advanced Mathematics (Allyn & Bacon, Newton, MA, 1966).

[2] E. Hausen-Tropper, A Topological Theory of Learning using Multi-Linked Variable Schemata, in: *Proc. 3rd Internat. Conf. on Principles of Knowledge Representation and Reasoning* (Cambridge, MA, October 1992).

[3] E. Hausen-Tropper, VMS: variable multi-linked schemata, a representation for learning, in: *Proc. National Conf. on Artificial Intelligence*, AAAI-92; Workshop on *Knowledge Representation Aspects of Knowledge Acquisition* (San Jose, CA, 1992).

[4] R.D. Hawkins and E.R. Kandel, Steps toward a cell-biological alphabet for elementary forms of learning, in: G. Lynch, J.L. McGaugh, N.M. Weinberger, eds., *Neurobiology of Learning and Memory*, (Guiford, 1984).

[5] D.O. Hebb, *The Organization of Behavior* (Wiley, New York, 1944).

[6] D.O. Hebb, *A Textbook of Psychology* (W.B. Saunders, 2nd ed., 1966).

[7] H. Hecaen and M.L. Albert, eds., *Human Neuropsychology* (Wiley, New York, 1978).

[8] E.R. Kandel and J.H. Schwartz, eds., *Principles of Neural Science.* (Elsevier, Amsterdam, 1981).

[9] M.M. Merzenich and J.H. Kaas, Principles of organization of sensory-perceptual systems in mammals, *Progr. Psychobiology Physiological Psychology* **9** (1980) 1–42.

[10] R.S. Michalski, Knowledge acquisition through conceptual clustering: a theoretical framework and an algorithm for partioning data into conjunctive concepts, *Policy Anal. Inform. Systems* **4** (3) (1980) 219–244.

[11] R.S. Michalski and R.E. Stepp, Automatic construction of classification: conceptual clustering versus numerical taxonomy, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-5** (4) (1983) 396–410.

[12] J.L. Phillips, *The Origins of Intellect: Piaget's Theory* (Freeman, New York, 1969).
[13] J. Piaget, *Judgement and Reasoning in the Child* (Harcourt, Brace and World; original french edition, 1924; 1928).
[14] J. Piaget, *The Psychology of Intelligence*, (Routledge & Kegan Paul, London, Original french edition, 1947; 1950).
[15] J. Piaget, *The Origins of Intelligence in Children*, (International Universities Press; original french edition, 1936; 1952).
[16] J. Piaget, Development and learning, *J. Res. Sci. Teaching* **2** (1964) 176–186.
[17] F.O. Schmitt, F.G. Worden, G. Adelman and S.G. Dennis, eds., *The Organization of the Cerebral Cortex* (MIT press, Cambridge, MA, 1981).
[18] L.R. Squire and N. Butters, *Neyuropsychology of Memory*, (Guilford, 1984).