7

Machine Learning of Physics Word Problems: A Preliminary Report

PATRICK SUPPES, MICHAEL BÖTTNER AND LIN LIANG

1 Introduction

In this article we continue the research on machine learning of natural language begun in Suppes et al. 1992, which is the first publication, although the research started in 1989. Until recently we have concentrated on a natural robotic language for elementary assembly actions (Suppes et al. 1995, Suppes et al. 1995). In the present article we turn to machine learning of physics word problems. The same basic axioms of learning are used for these rather different uses of language. In spite of the considerable interest in physics word problems by cognitive scientists (e.g., Larkin 1983), many of the persuasion that human beings are, above all, formal symbol processors, no sustained effort, as far as we know, has previously been made to study in detail what conceptual apparatus is needed to read such problems and produce as output the desired equations. Bobrow's system STUDENT (1964) dealt with high school algebra word problems, but had no machine learning component. The report of our first efforts to construct such a machine-learning program is given here. As will be evident, the results are certainly preliminary, but they do reflect our more extensive past experience with the machine learning of robotic language.

We emphasize that our machine-learning program, which has been applied to ten natural languages (see the references), assumes no prior knowledge of the target language. The only given knowledge is an alphabet and the fact that a word is an unbroken string of letters of the alphabet. It does have an internal language, essentially a language for physical equations, that is far removed from any given natural language,

Computing Natural Language
Atocha Aliseda, Rob van Glabbeek, and Dag Westerståhl, editors.

and in particular from English. Details of this internal language are given in Section 2. Key notions of our learning theory are working memory, long-term mempry, association, generalization, denotational value, and memory trace. These central concepts of our theory are explained in detail in Section 3, together with examples sketching their application to physics word problems. Here we remark on their intuitive origin.

The concept of generalization is widely used in psychological theories of learning. The principle of association goes back at least to Aristotle, and certainly was used extensively by eighteenth-century philosophers like Hume long before psychology had become an experimental science. The fundamental role of association as a basis for conditioning is thoroughly recognized in modern neuroscience and is essential to the experimental study of the neuronal activity of a variety of animals. For similar reasons its role is just as central to the learning theory of neural networks, now rapidly developing in many different directions. Our distinction about kinds of memory is standard in psychological studies of human memory, but the details of our machine-learning process are not necessarily faithful to human learning of language, and we make no claim that they are. On the other hand, our basic processes of association, generalization, specification and rule generation almost certainly have analogues in human learning, some better understood than others at the present time. In the general axioms formulated in this section we assume rather little about the specific language of the internal representation, although the examples that illustrate the axioms use the internal language described in section 2.

Another notion that is characteristic of our approach to learning is the concept of a denotational value. The denotational value of a word is the probability of whether that word has a denotation relative to the specific semantics of the problem domain. Not all words have a denotation relative to the given semantics. Which words have a denotation is largely determined by the internal language. In the robotic case our internal language has objects, properties, spatial relations, and actions. Therefore words not denoting anything that belongs to one of these categories are non-denoting. It is one of the purposes of learning to distinguish denoting words from nondenoting words. Before learning occurs a denotational value of 1 is assigned to every word. This value then gets reduced for non-denoting words during learning.

Section 4 summarizes our preliminary results and discusses some problems expected to arise. A formal statement of our learning axioms is given in the appendix.

2 Internal Language

As in the case of our work on robotic language, we concentrate on language learning and not on concept formation. This means that we endow the program with an internal language having a well-defined grammar and semantics. This internal language is not learned. It is used to interpret and learn the natural language,—in the present case English—, used to formulate physics word problems.

The internal language formulated here is for simple one-dimensional kinematic problems, but can easily be extended to cover other physics problems. The important point is that it is essentially an equational language, in which the reasoning is about equations, especially about the equations that express the data of a problem such as initial and final conditions, which may be formulated for a variety of physical quantities.

In the elementary problems which we have studied, the basic background assumptions are these. (i) Only kinematics of bodies is analyzed; no dynamics. (ii) Physical bodies are treated as point particles. (iii) Acceleration is always at a constant or uniform rate. (iv) The problems are all of one spatial dimension. (v) No derivatives of positions or velocity are introduced. Consequently the only relevant data or answers to questions are in terms of the following physical quantities: initial time t_0, initial position x_0, initial velocity $v(t_0)$; final time t_1, final position x_1, final velocity $v(t_1)$; elapsed distance Δx, elapsed time Δt; change in velocity Δv; acceleration a. So we can write in the internal language physical equations such as

$$v(t) = 10\ m/s, \qquad x_o = 5.1\ m,$$
$$t_1 = 60.2\ s, \qquad \Delta t = 5\ s.$$

where m is the symbol for *meters*, s for seconds, v for velocity, and t for time.

In contradistinction to the robotic language we have no category of objects. At the level of physics considered here all bodies are considered as point particles, and so we do not differentiate between properties of cars, trucks, tricycles and balls, because mass or weight is not a kinematical property. This is an important kind of abstraction used in physics, one which students must learn to do word problems efficiently.

Here is a very simple problem to illustrate our method of analysis. We directly give the representation in the internal language used to make semantic computations.

A	*car*	*accelerates*	*from*			$3.1\ m/s$
			$(t_0 = t$	&		$v(t) = 3.1\ m/s)$

to	6.9 m/s	in	5.0 s.
$(t_1 = t$ &	$v(t) = 6.9\ m/s)$	$(\Delta t = q$ &	$q = 5.0\ s)$

What is its acceleration?
$a(t) = ?$

The constant t_0 in the equation $t_0 = t$ is one of the semantic interpretations of *from*, special for physics word problems. The term 3.1 m/s in the equation $v(t) = 3.1\ m/s$ is the same in the internal representation. We use the units to determine the physical quantity v, but the time argument is left variable, to be determined by using the interpretation of *from* or *to*. Thus, the phrase *from* 3.1 m/s has the interpretation

$$(t_0 = t\ \&\ v(t) = 3.1\ m/s)$$

and by the logic of identity we then infer

$$v(t_0) = 3.1\ m/s.$$

The analysis of *to* 6.9 m/s is very similar, so that after the same sort of logical inference the interpretation is $v(t_1) = 6.9\ m/s$. The analysis of *in 5.0 s* uses the same setup, as can be seen from the analysis given. The computation for the question posed at the end of the problem goes along the same lines. The equational computations given above are trivial, but having a program that learns this special computational semantics and its associated grammar to solve this given class of problems is not.

Our non-lexical categories are

- W - for word problem (and start symbol),
- EC - for equation condition,
- E - for equation.

Our lexical categories are

- Q_V, Q_P, Q_T, Q_A - for physical quantity of velocity, position, time, and acceleration,
- R - for real number,
- U_V, U_P, U_T, U_A - for physical unit of velocity, time, and acceleration.

We have as terminal symbols real numbers and units $m, s, m/s, m/s^2$, and ?. No special category is assigned to the symbols v for velocity, x for position and t, t_0, t_1 for moments of time since for present purposes it is equations we are essentially working with.

The internal language is defined by the context-free grammar in Table 7. The strong compositionality of the internal equational language is transferred in learning to the denoting parts of the natural language — English in this case.

W	$\rightarrow$	EC
EC	$\rightarrow$	$EC\ EC/(E\ \&\ E)/E$
E	$\rightarrow$	$v(t_0) = Q_V/v(t_1) = Q_V/v(t) = Q_V/\Delta v = Q_V/$ $t_0 = t/t_1 = t/a(t) = Q_A$
E	$\rightarrow$	$x(t_0) = Q_P/x(t_1) = Q_P/q = Q_T/\Delta x = Q_P/$ $\Delta t = Q_T/\Delta t = q$
Q_V	$\rightarrow$	$R\ U_V/?$
Q_P	$\rightarrow$	$R\ U_P/?$
Q_T	$\rightarrow$	$R\ U_T/?$
Q_A	$\rightarrow$	$R\ U_A/?$
U_V	$\rightarrow$	$m/s, \ldots$
U_T	$\rightarrow$	$s, \ldots$
U_A	$\rightarrow$	$m/s^2, \ldots$

TABLE 7 Internal Grammar for Uniform Motion Problems

3 Theory

Basic Notions of our Theory Everything that is learned is stored in the learner's *memory*. It consists of two parts: a working memory to hold its content for the time period of a single trial and a long-term memory to store associations of words, denotational values, associations of grammatical forms and memory traces.

The long-term memory is not empty at the beginning but has stored in it an *internal language*. In the present study this internal language is stored in memory prior to learning and does not undergo any change during learning.

Whatever gets into the memory gets there by *association*. We use this concept to establish the connection between linguistic expressions and their meanings. Here, formally association is a binary relation between discourses, words and grammatical forms, on the one hand, and their respective counterparts in the internal language, on the other hand. In the present case of physics word problems, selected terms occurring in the natural-language statement of the problem are associated to equations in the internal language. For example, the preposition *from* is ordinarily correctly associated to an initial condition such as $t = t_0$ or $x = x_0$.

In our theory of learning we make a sharp distinction between denoting words and nondenoting words. Intuitively, only denoting words should acquire associations to terms or equations of the internal language. In the example used above the numbers together with their units are denoting. And so are the prepositions *from*, *to*, and *in*. All the other words ocurring in the word problem fall into the class of nondenoting words, the indefinite article *a*, the copula *is*, the possessive pronoun *its*, the substantive *car*, the verb *accelerates*. Especially for the case of nouns and verbs this may not look very natural from the point of view of conventional intuition. We emphasize however that we do not have in mind an absolute notion of denoting. What counts as a denoting word is determined by the internal language currently used in conjunction with a set A fixing the set of expressions available for association. Since the internal language varies from one domain of application to another so does the distinction of denoting versus nondenoting. In our current application, what counts as denoting is determined solely by the austere ontology of the equational language of physics.

By the *denotational value* of a word we understand the dynamically changing probability of that word having a denotation. If this value is 1 the word is denoting and if it is 0 the word is nondenoting. The purpose of this notion is to prevent nondenoting words from entering again and again into the probabilistic association procedure. We thereby exploit the fact that nondenoting words like, e.g., *the*, *a*, and *is* have a higher frequency of occurrence and should be learned more easily than denoting words, which have less frequent occurrences. Consequently, we set the initial denotational value to be 1 for all words, for we assume no prior knowledge of which words are denoting in a given language. Denotation learning follows a linear learning model:

$$(1)\quad d_{n+1}(a) = \begin{cases} (1-\theta)d_n(a) + \theta & \text{if } a \text{ occurs in trial } n \\ & \text{and is associated,} \\ (1-\theta)d_n(a) & \text{if } a \text{ occurs in trial } n \\ & \text{and is not associated,} \\ d_n(a) & \text{if } a \text{ does not occur in trial } n. \end{cases}$$

From various past experiments, we set the learning parameter $\Theta = 0.03$.

To show how the computation of denotational value works, let us consider further the associations given are $from \sim t_0 = t$. $accelerates \sim t_1 = t$. Let us further assume that at the end of this trial

$$\begin{aligned} d(accelerates) &= 0.900 \\ d(from) &= 0.950 \\ d(car) &= 0.700. \end{aligned}$$

On the next trial the sentence is

A truck accelerates from 2.5 m/s.

As a result, the association of *accelerates* is broken according to Axiom 2.6 i. Using $\theta = 0.03$, as we usually do, we now have $d(accelerates) = 0.873$, $d(from) = 0.9515$, $d(car) = 0.700$. After, let us say, three more occurrences of *accelerates* without any association being formed the denotational value would be further reduced to 0.620.

The dynamical computation of denotational value continues after initial learning even when no mistakes are being made. As a consequence high-frequency words have their denotational values approach to zero rather quickly. (From a formal point of view, it is useful to define a word as *nondenoting* if its asymptotic denotational value is zero, or, more realistically, below a certain threshold.)

The purpose of our principle of *generalization* is to generate grammatical forms. For example, the phrase *from 3.1 m/s* generalizes to the grammatical form $from\ R\ U_V$ where R is the category of real numbers and U_V the category of velocity units. Likewise the associated equation $(t = t_0\ \&\ v(t_0) = 3.1\ m/s)$ is generalized to the internal grammatical form $(t = t_0\ \&\ v(t_0) = R\ U_V)$.

When a generalization is made, the particular word association on which it is based is stored with it in long-term memory, as the *memory trace* justifying the generalization. The memory trace maps associations of grammatical forms to sets of word associations. The memory trace of a grammatical-form association thus keeps track of those word associations that gave rise to this particular grammatical-form association. If one of the word associations in the trace of a grammatical-form association is deleted, then so is this association.

The theory that underlies our learning program is given in terms of a system of axioms (for formal details see the appendix). The full set of axioms together with a detailed explanation of each axiom can be found in Suppes et al. 1996. We begin with a general formulation, which is then made more special and technical for learning the language of physics word problems.

Background Assumptions We state informally as background assumptions two essential aspects of any language learning device. First, how is the internal representation generated by the learner of an utterance heard or read, for example, for the first time. Second, at the other end of the comprehension process, so to speak, is that of generating an internal representation of a new utterance, but one that falls within the grammar and semantics already constructed by the learner.

Both of these processes ultimately require thorough formal analysis in any complete theory, but, as will become clear later, this analysis is not necessary for our present purpose. We give only a schematic formulation here.

1. *Association by contiguity.* When a learner is presented with a verbal stimulus that it cannot interpret then it associates the stimulus

		Production Rules	Grammatical-Form Associations		
W	$\rightarrow$	*a car accelerates EC.*	*a car accelerates EC.*	$\sim$	EC
EC	$\rightarrow$	$EC\ EC$	$EC\ EC'$	$\sim$	$EC\ EC'$
EC	$\rightarrow$	$(E\ E')$	$E\ E'$	$\sim$	$(E\ \&\ E')$
E	$\rightarrow$	Q_V	Q_V	$\sim$	$v(t) = Q_V$
E	$\rightarrow$	Q_T	Q_T	$\sim$	$\Delta t = Q_T$
Q_V	$\rightarrow$	$R\ U_V$	$R\ U_V$	$\sim$	$R\ U_V$
Q_T	$\rightarrow$	$R\ U_T$	$R\ U_T$	$\sim$	$R\ U_T$
E	$\rightarrow$	*from*	*from*	$\sim$	$t_0 = t$
E	$\rightarrow$	*to*	*to*	$\sim$	$t_1 = t$
E	$\rightarrow$	*in*	*in*	$\sim$	$\Delta t = t$

TABLE 8 Partial English Comprehension Grammar

to the single correct internal representation, whose structure will vary from one stimulus to another.

2. *Comprehension-and-response axiom.* If a learner is presented a verbal stimulus, then by using the associations and grammatical rules stored in long-term memory, the learner attempts to construct a semantic representation of the stimulus and respond accordingly.

Probabilistic Learning Algorithm For purposes of exposition we first describe the target state of learning, when some portion of English has already been learned. In a second step we describe how to reach this state from scratch.

If the learner has already learned enough English as to understand *A car accelerates from 3.1 m/s to 6.9 m/s in 5 s*, the learner's memory should contain (i) a semantically interpreted lexicon, (ii) a grammar of English, and (iii) a compositional semantics for that grammar like that shown in Table 8. With this memory the learner will be able to interpret the English word problem *A car accelerates from 3.1 m/s* by deriving its internal language translation $(t = t_0 \ \&\ v(t) = 3.1\ m/s)$. Given a more extended grammar than the one in Table 8 the derivation of the English word problem given initially is then completely straightforward.

In the following we shall describe how the memory can reach this state by learning from examples. We distinguish two cases: either no learning has occurred or some learning has occurred already. If no learning has occurred, the memory holds at least the internal language part. Unlike the robotic learning situation there is some knowledge about the English grammar already: the memory holds real numbers and physical units together with rules introducing the appropriate categories for them:

$$Q_V \rightarrow R\ U_V \quad U_V \rightarrow m/s, ..., \quad R \rightarrow 2.1, ...,$$
$$Q_T \rightarrow R\ U_T \quad U_T \rightarrow s, ...$$

Whenever a word problem given to the learner is not understood it is presented together with its internal language representation. A pair

(2) *A car accelerates from 3.1 m/s* $\sim (t = t_0\ \&\ v(t_0) = 3.1\ m/s)$

is formed from the English word problem and the learner's internal language counterpart.

Notice that our learning procedure does not start from scratch because it already has the real numbers and units occurring in English as part of the internal language. So the problem that remains to be solved is to find the English word to be associated to $t_0 = t$. Since there are 4 English words left, namely *a*, *car*, *accelerates*, and *from*, there are 4 possibilities to associate this equation. The learner probabilistically associates the words of the natural-language expression with the symbols of the internal-language expression. The probability that the learner associates *from* to $t_0 = t$ is only 1/4. Let us assume this indeed happens:

(3) $$from \sim t_0 = t.$$

By a principle of generalization (Axiom 1.2) the learner derives grammatical forms for both languages and derives the association

(4) $$a\ car\ accelerates\ E\ R\ U_V \sim E\ \&\ v(t) = R\ U_V.$$

The grammatical form will be stored in conjunction with those associations upon which the generalization was made.

By a principle of form association (Axiom 1.4) this association (4) will get broken down into smaller units like this:

(5) $$Q_V \sim v(t) = Q_V$$

In the internal equational language for the physics word problems we have the derivation

(6) $$E \rightarrow v(t) = Q_V.$$

From this and (5), we infer by a principle of rule generation (Axiom 1.3) the grammatical rule

$$E \rightarrow Q_V.$$

In a similar fashion, many other grammatical rules are generated:

(7)
$$\begin{aligned} EC &\rightarrow E \\ E &\rightarrow v(t) = Q_V \\ Q_V &\rightarrow R\ U_V \\ W &\rightarrow a\ car\ accelerates\ EC. \end{aligned}$$

By a principle of factorization (Axiom 2.2) we get from (4) and (5) together with grammatical rule (xviii) more general grammatical forms:

(8) $$a\ car\ accelerates\ E\ E' \sim (E\ \&\ E').$$

By a principle of filtering (Axiom 2.3) the set of grammatical forms and rules generated gets then reduced to a minimal set, e.g., (4) is replaced by the following form association:

(9) $$W \sim a\ car\ accelerates\ E\ E'.$$

Taking (3), (5) and (7) together, our memory will contain exactly what is needed to be able to understand part of the original word problem.

Consider now one of the cases with a wrong association hypothesized. An association that could arise with equal probability is

(10) $$car \sim t_0 = t.$$

By the principle of generalization (Axiom 1.2) we would now arrive at the association of the following grammatical forms:

(11) $$a\ E\ accelerates\ from\ Q_V \sim (E\ \&\ v(t_0) = Q_V).$$

Assume in the next trial the word problem *A car accelerates to 3.1 m/s* would be presented to the learner. Then the learner might generate the internal representation

$$(t_0 = t\ \&\ v(t) = 3.1\ m/s).$$

But this representation would be wrong since it assigns the car 3.1 m/s as its initial velocity rather than as its final velocity. Coercing the correct internal representation would result in breaking the association $car \sim t_0 = t$. The words *a*, *car*, *accelerate*, and *to* would reenter the sampling process with an equal probability to be associated to $t_1 = t$.

4 Some Results and Problems to be solved

In Figure 1 we show the mean denotational learning curve for the two nondenoting words *the* and *a* in our corpus. This curve is for an initial sample of 60 of our 105 training sentences based on 100 runs of 300 trials each. The computation for this learning curve is derived from the linear learning model defined in (1).

Note that by the end of the 300 trials, — halfway between 14 and 16 on the coordinates of the abscissa —, the denotational value of the two words is close to 0.1, and the denotational value of the denoting words remains close to 1.0, which was the initial denotational value, $d_1(w)$, for all words. A much more detailed treatment of the concept of denotational value is given in Suppes and Liang 1996.

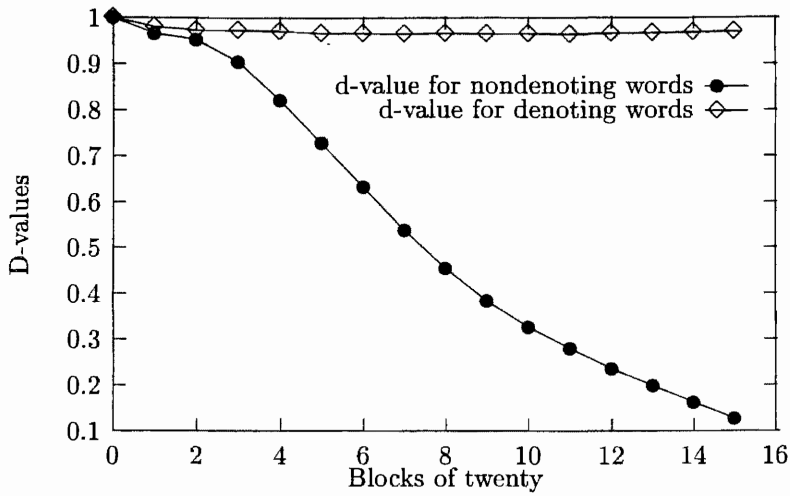

FIGURE 1 Mean learning curves for 60 English word problems.

We conclude with showing our equational analysis for two word problems that would require extension, but obvious ones, of the internal language and the English comprehension grammar given here. We do not discuss methods for actually solving these additional kinds of problems but computationally this is an easy task.

1. *Bob's car starts from rest and in 6.6 s has a velocity of 100 km/h. What is its average acceleration?*

   $(t_0 = t \ \& \ v(t) = 0)$
   $\Delta t = 6.6$ s
   $(t_1 = t \ \& \ v(t) = 100$ km/h$)$
   $a =?$

2. *Bob's car is now traveling east at 80 km/h and 1 km to the east, Susan's car is traveling west at 90 km/h on the same road. When will they pass each other?*

   $(t_0 = t \ \& \ v_b(t) = 80 \ km/h \ \& \ x_s(t) - x_b(t) = 1 \ km \ \& \ v_s(t) = -90 \ km/h)$
   $(t_1 = t \ \& \ x_b(t) = x_s(t) \ \& \ t_1 - t_0 =?)$

There is a big difference between our robotic learning experiment and the physics learning experiment. In the robotic experiment the natural language used was simple, consisting mainly of commands to manipulate an object of a limited environment in the robot's perceptual field. On the other hand, the notions involved (objects, properties, spatial relations, actions) had largely been left undetermined. In the case of physics word problems just the opposite is the case: all the notions involved can be given a rigorous definition, but the language of word problems is much richer in variety of expression with respect to contextual interpretation and anaphora.

Our learning program certainly does not reflect the approach of a human learner. For which human learner would it make sense after all to learn a language from physics text books? However, we would like to draw attention to a phenomenon that has been observed with students familiar with the physical theory underlying the word problem but not deeply familiar with the language in which the problem is presented: they generally have not much difficulty in understanding and solving these problems. For the same reason we expect our project to succeed: in that respect we think our program resembles the student who is in command of physics but not in good command of the language to be learned. We therefore conjecture that their language learning does not depend on compositionality but resembles more the learning of a fixed list of words and phrases. These phrases refer to certain qualitative notions like distance, duration, speed, acceleration. We therefore are currently developing a qualitative semantics of physical processes that fits more closely the structure of natural language than does the equational language of physics.

References

Larkin, Jill. 1983. The Role of Problem Representation in Physics. In *Mental Models*, ed. Dedre Gentner and Albert L. Stevens. 75–98. Lawrence Erlbaum.

Suppes, Patrick, Michael Böttner, and Lin Liang. 1995. Comprehension grammars generated from machine learning of natural languages. *Machine Learning* 19:133–152.

Suppes, Patrick, Michael Böttner, and Lin Liang. 1996. Machine Learning Comprehension Grammars for Ten Languages. *Computational Linguistics* 22(3):329–350.

Suppes, Patrick, Michael Böttner, Lin Liang, and Ray Ravaglia. 1995. Machine Learning of Natural Language: Problems and Prospects. In *WOFCAI 95. Proceedings of the Second World Conference on the Fundamentals of Artificial Intelligence*, ed. Michael de Glas and Z. Pawlak, 511–525. Paris. Angkor.

Suppes, Patrick, and Lin Liang. 1996. Probabilistic association and denotation in machine learning of natural language. In *Computational Learning and Probabilistic Reasoning*, ed. A. Gammerman. 87–100. John Wiley & Sons Ltd.

Suppes, Patrick, Lin Liang, and Michael Böttner. 1992. Complexity issues in robotic machine learning of natural language. In *Modeling Complexity Phenomena*, ed. Lui Lam and Vladimir Naroditsky. 102–127. Springer.

Appendix

Axioms of Learning

1. **Computations using Working Memory**

1.1 ***Probabilistic Association.*** On any trial, let s be associated to σ, let a be in the set of words of s not associated to any internal expression of σ, and let A be the set of expressions of the internal language made available for association and let α be in A but not currently associated with any word of s. Then pairs (a, α) are sampled, possibly using the current denotational value, and associated, i.e. $a \sim \alpha$.

1.2 ***Form Generalization.*** If $g(g'_i) \sim \gamma(\gamma'_i)$, $g'_i \sim \gamma'_i$, and γ' is derivable from X, then $g(X_i) \sim \gamma(X_i)$, where i is the index of occurrence.

1.3 ***Grammar – Rule Generation.*** If $g \sim \gamma$ and γ is derivable from X, then $X \rightarrow g$.

1.4 ***Form Association.*** If $g(g') \sim \gamma(\gamma')$ and g' and γ' have the corresponding indexed categories, then $g' \sim \gamma'$.

1.5 ***Form Specification.*** If $g(X_i) \sim \gamma(X_i)$, $g' \sim \gamma'$, and γ is derivable from X, then $g(g'_i) \sim \gamma(\gamma'_i)$.

1.6 ***Content Deletion.*** The content of working memory is deleted at the end of each trial.

2. **Changes in State of Long-term Memory**

2.1 ***Denotational Value Computation.*** If at the end of trial n a word a in the presented verbal stimulus is associated with some internal expression α, then $d(a)$, the denotational value of a increases and if a is not so associated $d(a)$ decreases. Moreover, if a word a does not occur on a trial, then $d(a)$ stays the same unless the association of a to an internal expression α is broken on the trial, in which case $d(a)$ decreases.

2.2 ***Form Factorization.*** If $g \sim \gamma$ and g' is a substring of g that is already in long-term memory and g' and γ' are derivable from X, then g and γ are reduced to $g(X)$ and $\gamma(X)$. Also $g(X) \sim \gamma(X)$ is stored in long-term memory, as is the corresponding grammatical rule generated by Axiom 1.4.

2.3 ***Form Filtering.*** Associations and grammatical rules are removed from long-term memory at any time if they can be generated.

2.4 ***Congruence Computation.*** If w is a substring of g, w' is a substring of g' and they are such that

i. $g \sim \gamma$ and $g' \sim \gamma$,

ii. g' differs from g only in the occurrence of w' in place of w,

iii. w and w' contain no words of high denotational value,

then $w' \approx w$ and the congruence is stored in long-term memory.

2.5 ***Formation of Memory Trace.*** The first time a form generalization, grammatical rule or congruence is formed, the word associations on which the generalization, grammatical rule or congruence is based are stored with it in long-term memory.

2.6 ***Deletion of Associations.***

i. When a word in a sentence is given a new association, any prior association of that word is deleted from long-term memory.

ii. If $a \sim \alpha$ at the beginning of a trial, a appears in the utterance s given on that trial but α does not appear in the internal representation σ of s, then the association $a \sim \alpha$ is deleted from long-term memory.

iii. If no internal representation is generated from the occurrence of a sentence s, σ is then given as the correct internal representation, and there are several words in s associated to an internal expression α of σ such that the number of occurrences of these words is greater than the number of occurrences of α in σ, then these associations are deleted.

2.7 ***Deletion of Form Association or Grammatical Rule.*** If $a \sim \alpha$ is deleted, then any form generalization, grammatical rule or congruence for which $a \sim \alpha$ is a memory trace is also deleted from long-term memory.