An Engineering Model of Learning

Michael H.W. Hoffmann
Vice-Head, Dept. of Microwave Techniques, University of Ulm,
Albert-Einstein-Allee 41, D-89069 Ulm, Germany, Michael.Hoffmann@ieee.org

Abstract - Learning theories that are favored by psychologists and by industry for education of adult learners turn out to be too simplistic for application to engineering education. An integrated learning model that is taking into account recent results from cognitive psychology, from neurophysiology, and from information processing appears not to be available. Therefore, this paper is aimed at designing a novel, interdisciplinary model of learning from an engineer's point of view. Learning will be described as an adaptive and nested feedback control process that comprises different levels of learning, as reacting automatically to recognized situations, training of skillfully handling decisions, or handling abstract ideas. Thus it might be explained why learning to understand abstract ideas takes considerably more time than learning to handle situations by rote. From that, conclusions might be drawn concerning mediating knowledge in classroom, or on designing complete curricula for engineering education.

Index Terms – Curriculum design, instructional design, learning theories, taxonomy of learning.

INTRODUCTION

While learning theories are being applied more and more in continuing education, their acceptance in engineering education is rather inert. Many lecturers in engineering feel that existing learning models do not explain learning of complex relations satisfactorily. They are instructions for training rather than learning models (see for instance [1]), or they are pieces of advice from experienced instructors to younger colleagues (e.g. [2], [3]).

The mentioned models undisputedly describe main steps of the learning process, but they are too coarse to understand different levels of learning, or the difference between learning purely factual knowledge (basic knowledge of facts), conceptual knowledge (interrelations between facts), procedural knowledge (how to apply methods for using factual and conceptional knowledge), and metacognitive knowledge (knowledge of one’s own cognition). Neither do they explain satisfyingly the flow of information during the processes of remembering, nor of understanding interrelations, of applying knowledge, of analyzing problems, of evaluating solutions, or of creating new solutions.

The terms used in the last paragraph are taken from Anderson’s and Krathwohl’s revision of Bloom’s taxonomy of educational objectives [4], or of “Bloom’s taxonomy” as it will be called for the rest of this paper.

Many of the known learning theories definitely present highly interesting ideas on the adaptation of information, see for example the theories of Skinner [5], a representative of behaviorist theories, or of Gagné [2] and co-workers as representatives of cognitive theories, or of Bandura [6], and Lewin [7] being representatives of constructivist theories, or of Piaget and his school [8] being representatives of both constructivist and cognitive theories, just to mention a few. (A concise overview on the major learning theory schools might be found in [8]).

However, only few theories are available that combine ideas from cognitive neurology with those from information processing. Therefore, it is attempted to find a novel interdisciplinary model of learning from an engineer’s point of view, where the flow of information during learning is described.

METHODS

To find a new model and – based upon it – a new theory of learning, a combination of a top-down-design method and of a bottom-up-design method of complex information processing systems is applied.

Starting with a definition of the term “learning”, a block circuit diagram is developed that attempts to describe learning by functions that are necessary to be performed during the process of learning.

If possible, then the functional blocks are tested against modern experimental findings, e.g. from neuroimaging methods [10] as for instance functional magnetic resonance imaging (fMRI), or positron emission tomography (PET), or other well-accepted experimental methods.

The signal-flow oriented method is a method that is not only successfully applied in engineering sciences, but also in cognitive neurosciences and neuropsychology. Baddeley’s and Hitch’s [11] model of the working memory is a good example of this method.

Though published findings using neuroimaging methods are strongly supporting the plausibility of the model, its validity has not yet been rigorously tested by experiments, since this will need cooperation of scientists of several disciplines.

THE MODEL

I. An overview

The development of the model starts with a definition of learning. Following Shuell [12], “nearly all conceptions of learning have involved – either explicitly or implicitly – three
criteria: a change in an individual’s behavior or ability to do something, a stipulation that this change must result from some sort of practice or experience, and a stipulation that the change is an enduring one.”

Since an individual is not very likely to change “behavior”, if there is no good reason for doing so, a criterion of comparison and assessment is added. Therefore, in this paper, the term “learning” is seen as a process in which the assessment of a given set of information and a list of potential consequences by that assessment are changed enduringly as the result of comparative observation or comparison of actual experience with former experience.

The process of learning implies the detection of new or uncommon pieces of perceived information: If there is nothing detected to change a learner’s attitude, then the learning process is not initiated. Therefore this process of detection or recognition is of major importance for the model.

Experiments using functional magnetic resonance imaging (fMRI) show that recognition of information is a multiple step procedure. It will be demonstrated that this also implies learning to be a multi-step-procedure. As a result, learning might be described as a process where many similar cycles are run through and where each cycle is modeled as in Fig. 1.

FIGURE 1
DIAGRAM OF A LEARNING CYCLE ON A SPECIFIC LEVEL OF LEARNING.

At first view, parts of the model look very similar to some existing models. The difference might be recognized when analyzing the procedures in the diagram and their structures. This will be done in more detail in a section below.

However, for getting an overview, the diagram might be explained by applying it to learning a topic that is well known to engineers. It is learning to handle polynomial expressions in mathematics.

Here, the learner must first get to know the definition of a polynomial. He or she acquires factual knowledge.

Next, the learner must recognize that polynomials mostly appear in a certain context, for example as a certain type of functions, or as a placeholder for an algebraic expression: the learner is acquiring conceptual knowledge.

In a third step, addition and multiplication of polynomials are trained time and again to memorize the complete process as a template for further usage, which allows for more or less automatic adaptation to polynomials of any degree. This is acquiring procedural knowledge. The learner might then revert to find the n-th power of polynomials and to analogously repeat the cycle.

With increasing experience, the learner might find some rules for the usage of polynomials. He or she might find out that polynomials might obey rules that are similar to the rules integer numbers obey: The learner is acquiring canonical knowledge. This opens the way for learning new facts by analogy, for example how to divide one polynomial by another polynomial: a cycle on a higher level of learning would be initiated that would be performed best, if the learner was successful on the lower level.

Instead of leaving a cycle on a certain level of learning after acquiring procedural or canonical knowledge, the learner could think about different scenarios, for instance what would happen if the degree of polynomials would tend to infinity. The acquired strategic knowledge would help the learner to avoid mistakes and to master certain aspects of infinite series. He or she could then revert to another cycle of learning, either on the same level, or on a higher level.

On relatively high levels of learning, the process of strategic learning could even be used to think about what would happen, if procedural knowledge would be applied in a different context, or if rules would be broken. This opens the way to creative thinking.

While the diagram shown in Fig. 1 is demonstrating some main ideas and an overview of the new model, some other aspects might only be seen by a closer analysis. This applies particularly to the validation of some details by neurological methods. Therefore, in the next sections, the subprocesses of learning are analyzed in closer detail.

II. The basic mechanism of recognition

In order to compare given information with internally binding rules or experience, it is necessary to compare input information with known pieces of information that are recalled from memory. These known pieces will be called patterns. Since these must be available very quickly and at any time, it is sensible to assume that their storage is organized differently from other data stored in long-term memory. This assumption is supported by findings from Yonelinas [13], who detected two different mechanisms for recognition in the episodic memory: one that measures “familiarity” (which is pattern recognition), and one that recollects data.

It is obvious that comparing perceived information with stored patterns can only work successfully if the list of patterns is not too large. This is confirmed by researchers in the field of speech recognition [14], [15], who found out that the process of recognizing spoken language is subdivided into a multi-step process of pattern recognition, which allows for small pattern lists. Indeed, fMRI experiments show that for different tasks of recognition different areas in the brain are used.

Therefore, in a very early stage of signal processing, information is coarsely classified. Thus, only a limited number of patterns must be compared to input information. Later,
information is classified finer and finer. This does not only apply to auditory information, but also to most other types of information, and in particular to visual information [16]. Pattern recognition does not result in a “recognized” – “not recognized” decision. Rather, a list of more or less fitting patterns is used for further decisions, as it is seen by experiments in speech recognition.

Thus, processing of information on a relatively basic level might be modeled as follows: sensory perception, be it visual, or auditory, or other, is compared in a comparator to a specialized list of patterns. From those patterns that are similar to the perceived information, the one with the largest similarity is identified as the most probable one.

From that, it is concluded that recognition in general is a process that consists of many sub-processes, and where each sub-process needs its own pattern store. Furthermore, it is seen that recognition works on different levels: Some patterns are only identified, if others have been identified before.

**III. Learning to recognize new basic patterns**

Recognition can only work successfully, if adequate patterns are available for comparison. Some basic patterns are certainly inherent to all human beings. But how are novel patterns found and placed at disposal?

In a few cases, particular events might be related to extreme emotions like extreme joy, or pain, or sadness. These events are immediately transferred for storage, and a copy of this piece of information is used as a pattern. Most other cases are treated differently, since classifying pieces of information as being worth storing them is not so clear cut. In these cases, a single occurrence of an event or of another piece of information is not enough to recognize their importance.

However, if a particular piece of information is repeatedly input to a specific comparator that delivers a “not known” decision, and if at the same time attention or emotion is linked with the repeated occurrence of it, then this piece of information has come into notice, which is stated by fMRI experiments [17]. Once the decision is made that a particular piece of information might serve as a new pattern, it is stored in the pattern store.

Apart from extreme situations, the process of getting to know a new pattern needs to be a repeated occurrence of a particular piece of information such that it can be distinguished from other information, and that a measure of importance is linked to it in order to classify it as being worth storing it. Therefore, there must be a functional block that detects information that is not (yet) stored in the pattern store and that decides on its importance. Since such a decision can only be made after comparison to known patterns, input information must be buffered to make it available again at a later point in time.

Not all stored patterns are constantly used by an appropriate comparator. This is proven by the fact that people might reload pattern lists when switching from one spoken language to another one [18]. It is therefore concluded that information available in the pattern store is buffered in a re-loadable buffer store before being used by the comparator.

**IV. Getting familiar with complex patterns**

In the neuropsychological literature, at least two different types of stores are identified in the human brain. These are short-term stores and long-term stores [19], [20], the former forgetting in intervals measured in some hundred milliseconds to some ten seconds, the latter forgetting in much larger intervals of time, up to decades. It is supposed that a process of interconnecting nerve-cells is responsible, or at least necessary for the formation of the long-term memory. This process is called “plasticity” in the literature of neurosciences. If that process is not consolidated by repeated recall, then it might be suspended, or even aborted.

The buffer-stores in the model following Figure 2 must be capable of adapting relatively quickly to changes of input information. Therefore, they must be implemented as short-term stores. In contrast to the buffers, the basic pattern store must be a long-term store. Therefore, repetition on a longer time-scale is necessary for storing patterns in the pattern store.

Therefore, recognizing new patterns and storing them might be modeled as shown in Figure 2. There, input information is buffered before feeding it to the comparator. The same happens to information stemming from the basic pattern store. The comparator delivers its decision and a measure for reliability. Buffered input information together with information about “importance” and frequency of appearance, and buffered information about patterns are used to detect unknown patterns that are worth storing them. The latter are fed back to the pattern store, where they are now available for comparison: New patterns are learned in a feedback loop!
If recognition concerns patterns that are composed of a series of simpler patterns, then the properties of the buffers could prevent successful recognition. The reason for this is that in that series of patterns, one after the other has to be processed. Meanwhile, the input buffer might have lost part of its information, since it is implemented as a short-term store.

The human brain has found at least two solutions to tackle that problem. These are two feedback-loops that were first hypothesized to exist by Baddeley and Hitch [11]. They are called “phonological loop” and “visuospatial sketchpad”. Both loops might be understood as loops including a short-term store, and a feedback-connection for refreshing the content of the store. Roughly speaking, in the visuospatial sketchpad sequences of information are stored that one might imagine with closed eyes; in the phonological loop sequences of information are stored that one might imagine as being spoken in one’s mind. Either of these loops might work as a replacement of the input buffer in the feedback loop for learning new complex patterns.

Following Baddeley and Hitch, the activities of both loops are controlled by a device they are calling “central executive” and that performs similar tasks as Shiffrin’s and Schneider’s “attention director” [21].

There are notable consequences from that assumption. Both of the loops, visuospatial sketchpad and phonological loop work best, if full attention is given to them. Otherwise, they are likely to be disturbed and to forget. In other words: recognizing complex information needs attention and sufficient time to work efficiently.

Thus, learning new complex patterns is not different in principle from learning basic patterns, though the requirements for awareness and for time are different as compared to learning basic patterns. What is common to both is that before memorizing a new pattern, repetition is necessary, as well as classification as being important. Also the consequences formulated in section III are the same.

In terms of learning taxonomy, learning new patterns might be identified as learning facts. The content of the long-term pattern stores might be interpreted as factual knowledge in the sense of Bloom’s taxonomy.

Cognitive neuroscience distinguishes between conscious (declarative, explicit) and unconscious (nondeclarative, implicit) memory [22], [23]. Applied to the new model, this means that pattern stores that were “filled” without awareness would belong to implicit memory, while pattern stores that got their content by usage of the phonological loop or the visuospatial sketchpad belong to explicit memory.

V. Development of conceptual knowledge

Identification of a piece of information as being known would be useless if there would not be some consequences. If for example an object is identified to move directly towards an observer, then knowledge of this fact would not be useful if the observer could not recognize potential consequences of that movement, e.g. that the object might hit the observer.

It is thus necessary to recognize, whether certain patterns could appear in combination, either at the same time or in a relatively close relation in time. I.e.: it is necessary to recognize a conceptional event.

In the above given example, it would be necessary to recognize that certain types of movements are combined with the experience of pain. Therefore, a second type of a learning system must be active in the human mind, and that recognizes whether known patterns are linked in a given context.

As in the process of learning to recognize new patterns, it is first necessary for this new learning system to recognize that two or more patterns appear in pairs (or n-tuples). Therefore, it is assumed that this is done in a system that is very similar to that described by Figure 2. However, instead of simple patterns, pairs or n-tuples of patterns are stored, from which it is known that they occur in combination. Unlike the system for simple pattern recognition, the new system compares the buffered pairs or n-tuples, the conceptional events, with combinations of patterns that have been recently detected by the simple systems for pattern recognition. The rest of the second system is acting exactly like the (first) system for pattern recognition.

This means in particular that – like in the case of simple pattern recognition – a novel conceptional event could be learned, provided the novel combination would occur repeatedly and with a sufficiently large measure of attention assigned to it.

The memory, where information about links between patterns is stored, must be different from the original pattern stores, since the content is different. This does not mean that the type of memory had to be different, only the individual store must be different.

A hint (though not strict evidence) to the correctness of these ideas is given by Hugdahl et al. [24], where listening to words with attention activates additional brain regions as compared to passive listening to words.

What is stored in the memory for conceptional events might be seen as conceptual knowledge in the sense of Bloom’s taxonomy.

VI. Evolution of procedural knowledge

Reverting to the example of an object that is moving directly towards an observer, it is obvious that factual and conceptual knowledge would be useless, if no action could be initiated in order to influence the situation. Therefore, a third learning system must be available that places one or several procedures at the disposal of the learner in order to give him/her an opportunity for (re-) action. In the given example, this could be performance of a movement to avoid collision.

More complex situations would need more complex actions. Thus, a list of actions would be a good response to recognition of an event.

These actions could be found either by copying actions from others or by own experience. I.e.: first an action must be recognized, then it must be assessed as successful and worth storing, and then it must be stored. This is again a learning process that uses the same steps of information processing as in the two systems discussed before. Therefore, the structure from Figure 2 could be re-used while replacing the pattern
store by an action store storing procedural lists, and where a list of actions to be recognized is taken as input.

Thus, learning by imitating finds its explanation: if an action is undertaken that was copied from others, and if this action was successful, it will be stored in one's own action store. In case there is no list of actions stored, own experience must serve as source of an action list. Own actions include “doing nothing”, which might lead to an uncomfortable experience as well as to good experience.

Again, hints supporting the correctness of these ideas are given by fMRI-studies, where Rowe et al. [25] report on context-specific changes in effective connectivity to motor or visual cortex during the selection of action or color.

What is stored in the action store might be seen as *procedural knowledge* in the sense of Bloom’s taxonomy, though with a somewhat restricted meaning. Again, this type of knowledge might be distinguished into knowledge stored in implicit or explicit memory.

## VII. Detection of rules

Many actions stored in procedural knowledge obey similar rules. A quite intriguing example is the application of such rules when using language. A native speaker uses these rules most of the time without consciously recalling them, while a non-native speaker with only little command of the language consciously retrieves these rules from memory. I.e.: there must be a store for such rules.

Abstractly speaking, detection of rules is nothing else than finding a common pattern behind a set of patterns. It is thus obvious that again an information processing system as shown in Figure 2 could be used for that task. This time, the pattern store must be replaced by a store containing rules. Input to this system is a list of patterns.

What is stored in the memory for rules might be named *canonical knowledge*. It is *not* strictly coinciding with one of the knowledge types in the sense of Bloom’s taxonomy.

## VIII. Foundations of creativity

Finding own actions as a reaction to given events is the basis of being more successful than others. There is not only the possibility of just doing something to find new successful actions. There is also the alternative to run through virtual scenarios in order to find out what would happen if something would be done.

Such a process must include activation of the visuospatial sketchpad and of the phonological loop. Moreover, it must activate access to episodic memory, where lists of personally experienced events are stored, and to canonical memory, where rules are stored. To handle such scenarios, an information processing structure might be used, that is close to the system shown in Figure 2. This time, the pattern store is replaced by a store for *strategies*, which might be seen as complex actions. Input to the system would be virtual episodes delivered by episodic memory or by action stores from procedural memory.

However, there are also some differences to the cases before. Firstly, the comparator must act dynamically following rules from the memory for rules.

Secondly, controlled by the central executive, and managed by the control unit in the feedback loop for learning new patterns (see Figure 2), it must be possible to suppress or replace part of the input episodes, and to break some of the rules. As a result, new strategies are discovered that might be learned (stored), if they appear interesting.

What is stored in the memory for strategies might be called *strategic knowledge*.

## CONSEQUENCES AND CONCLUSIONS

If the above described engineering model of learning is accepted as a hypothesis for explaining learning processes, then it has consequences for education in general, and for engineering education in particular.

### I. Learning is a multiple-step process on different levels

Learning is a process that in general consists of multiple subprocesses of recognition, of assessing and detection. Simple learning processes might be performed automatically. In contrast, *complex learning processes* involve subsystems such as the visuospatial sketchpad or the phonological loop, which require *attention* for efficient working.

The process of learning is interrupted if one or more of the involved subsystems are interrupted. E.g.: If the basis of factual knowledge is not broad enough, then conceptual knowledge will have gaps. If conceptual knowledge is not broad enough, procedural knowledge might not be applied correctly. Then rules cannot be detected, and strategies cannot be developed.

All types of knowledge might occur on different levels. Complex events for instance might be the factual basis of more complex conceptions and procedures. If the underlying basis has gaps, then some complex facts cannot be classified correctly, which hampers building up complex conceptual knowledge etc.

### II. Learning success needs motivation and repetition

Two well-known rules are confirmed: learning needs motivation and learning needs repetition.

The first rule is explained by the fact that learning needs attention or it must be linked with feelings. Therefore, the process of learning is supported by motivation.

The second rule is explained as follows: If a pattern, an event, a procedure, a rule or a strategy is not repeated several times, then it is not detected as something worth learning. Therefore, repetition is a must.

There are some reports where repetition allegedly would not improve the efficiency of learning. However, closer analysis of such reports shows that a learning process in these cases simply had not been initiated! Either there were not enough repetitions, or complex patterns were presented too briefly in order to be recognized.

### III. Learning needs time and thought

Each individual subprocess of learning needs time, first because repetition is necessary, second since forwarding...
information into long-term store needs a certain amount of time. The more subsystems of learning are involved in a learning process, the longer time it will take to learn.

Conceptual, canonical, and strategic knowledge on a higher level supposes active comparison of patterns, of procedures, and of real and imaginary episodes. This needs a considerable amount of time and of thought.

   If the processes of canonical and strategic learning are initiated and prematurely aborted, then the concerning types of knowledge will not be consolidated: they will disappear from memory. Consequently, part of the learning work done before is wasted.

IV. Learning by rote is not enough for acting skillfully

Facts might be learned by rote. However, if a fact to be learned is a complex composition of patterns, events, procedures, rules, or strategies, then the latter must be safely known before. Otherwise, they will or they will only partly be forwarded to the next learning subsystem. Procedures might also be learned by rote. However, if conceptual knowledge on the involved facts is not present, then actions are not taken skillfully.

   Acting skillfully means that procedural knowledge is not only accompanied by broad conceptual knowledge, but there must also be knowledge of underlying rules, and there must be a portfolio of strategies that could be applied, if a situation changes.

   This in turn demands that many levels of recognition and learning have been performed beforehand successfully. This means in particular that the learning process for skillfully acting takes considerably more time than that for learning by rote.

V. Curricula might be optimized for learning success

In order to educate students optimally, enough time must be provided in the curricula for consequently building-up knowledge of all types, and to network them.

   If, for example, time would be “saved” by cutting down mediating conceptual knowledge, then procedural knowledge could only be applied without deeper comprehension, which might be enough for basic education.

   However, if after a certain period of time education were to be improved, then much more time must be spent to build up new networked knowledge, which never had a chance to be consolidated due to missing prerequisites!

   Experienced lecturers know this effect: much of the knowledge that was supposed to be learned in a bachelor-course is no longer available in the beginning of a subsequent master-course. This might be due to “tightening” previous courses, and thus prematurely aborting the processes of canonical and strategic learning.

   Give students time to learn! In particular, give students time to build up canonical and strategic knowledge!

REFERENCES


