

LANGUAGE LEARNING AS PROBLEM SOLVING

Modelling logical aspects of inductive learning to generate sentences in French by man and machine

Michael ZOCK
Gil FRANCOPOULO
Abdellatif LAROUÏ

LIMSI, BP 30, 91406 Orsay - France

Abstract:

We present here a system under development, the present goals of which are to assist (a) students in inductively learning a set of rules to generate sentences in French, and (b) psychologists in gathering data on natural language learning.

Instead of claiming an all-encompassing model or theory, we prefer to elaborate a tool, which is general and flexible enough to permit the testing of various theories. By controlling parameters such as initial knowledge, the nature and order of the data, we can empirically determine how each parameter affects the efficiency of learning. Our ultimate goal is the modelling of human learning by machine.

Learning is viewed as problem-solving, i.e. as the creation and reduction of a search-space. By integrating the student into the process, that is, by encouraging him to ask an expert (the system) certain kinds of questions, like: can one say x ? how does one say x ? why does one say x ?, we can enhance not only the efficiency of the learning, but also our understanding of the underlying processes. By having a trace of the whole dialogue (what questions have been asked at what time), we should be able to infer the student's learning strategies.

1 THE PROBLEM OF LEARNING A LANGUAGE:

Language learning can be viewed as a special case of problem solving in which the learner tries to build and intelligently explore a hypothetical search space. If this view is correct, then two sets of questions arise immediately. On one hand one may want to know:

- what the nature of this search space is (what are the variables?),
- how it is built (incremental learning: local vs global view),
- how it is explored (strategies: intelligent, opportunistic vs systematic search).

On the other hand, one may want to investigate how (i) the knowledge at the outset and (ii) the ordering of the data will affect the building and the searching of the space. Typically one does not learn from scratch, nor is it likely that one encounters either well-ordered data, or a complete set of examples: natural learning is incremental.

Obviously, these facts imply that:

* initial knowledge, in particular, knowledge of other languages may bias the kind of variables (attributes or hypotheses) considered, i.e., included in the search space;

* the order of the data (the examples encountered by the student) may determine what rules are likely to be inferred at what moment, and finally

* rules are inferred from incomplete data (incremental learning). Furthermore, the same data may be characterized in different ways. That is, several equivalent descriptions may be inferred from the same data set. Which of these descriptions turns out to be the most adequate generally cannot be established until one knows the complete data set. Thus, rules may have to be revised in the light of new evidence. Consequently, errors are not only unavoidable parts of the learning process, but also an indispensable source of information for the learner.

2 THE PROBLEM OF TEACHING HOW TO LEARN:

As we have shown, learning can be seen as searching. Actually, teaching, as well as learning, can be conceived of as problem solving or reasoning in an information-exchange environment. There is a sender, a goal, a message and a receiver. The SENDER may be a native speaker, a teacher, a parent, a book or a computer. The GOAL is the task or performance (output). In our case it is knowledge of how to produce sentences in French. The MESSAGE is the input to the learning component: examples from which the rules have to be inferred (1). The RECEIVER or learner can be any system, natural or artificial, capable of perceiving, memorizing and analyzing a set of data and drawing the necessary conclusions: a child, a student, or a computer program (2).

Learning occurs in various settings. Depending on the order of the examples and the control of the information flow, we speak of natural, experimental, or institutional settings. Natural learning is characterized by the absence of a clearly defined learning objective (3), by noisy and heterogeneous material, and by unordered examples. The underlying regularities are thus multiple, diffuse, and hard to perceive. Experimental learning and teaching, on the other hand, have a learning objective, the material is error-free, homogeneous and coherently ordered according to some point of view (learner or teacher). Whereas experimental learning can be characterized by the following sequence: (i) encountering the data (ii) analysis, (iii) building and testing of hypothesis, (iv) feedback and (v) proof or demonstration of the theory, traditional teaching goes through the following stages: (i) exposition, (ii) practice, (iii) testing and (iv) evaluation. This can be schematized as follows:

Teacher: sets the task and presents the learning material;

Student: analyzes the data;

Teacher: provides a set of examples;

Student: practices;

Teacher: asks questions to test the gained knowledge;

Student: answers the questions;

Teacher: evaluates the answers, provides feedback (explanations) and organizes future data as a function of actual performance

Student: integrates the feedback into the knowledge base and corrects misconceptions;

As one can see, the information flow here is entirely teacher-controlled. He is the one who sets the task, and provides the examples and the feedback. Consequently, the teacher decides the nature and the order of the material to be learned.

There are two major shortcomings in this approach. Not knowing what information is needed by the learner, the teacher may present the wrong data. More importantly, the student is only loosely integrated in the learning process. Instead of being active, generating and testing plausible hypotheses (discovery learning), he reacts to questions. Thus, it may happen that the student perceives his task as the learning of the material rather than the learning of the underlying principles.

Ignorance of what or how to learn may result in (i) learning the unintended, (ii) poor problem-solving skills or (iii) little transfer. As long as the learner does not go beyond the information given (the concrete word level), he cannot transfer the gained knowledge to similar situations, because the perception of similarity presupposes abstraction.

Given these criticisms, it would be useful to have a system which has the qualities mentioned above without having the drawbacks. A good learning environment should be both flexible and constraining enough:

* to allow for simulation of real communication, that is to say, to provide a setting where both participants can take the initiative and control the information-flow,

* to ensure the learning of the appropriate material (i.e., what to learn) as well as the necessary problem-solving skills (the methods, i.e., how to learn).

A computer program could provide such an environment. It would offer different kinds of information (see below: trace-function), while answering the student's questions as he goes along generating and testing different sorts of hypotheses.

3 THE COGNITIVE ENGINEER'S TASK:

to provide the user a friendly interface

We will describe here a system under development, whose major goals are:

* to provide an environment which allows communication between a learner (student) and an expert (in our case the system);

* to simulate the information-processing aspect of natural learning, i.e., the inductive learning of grammatical rules to generate sentences in French.

* to allow teachers and psychologists to test various theories.

The system we have in mind is designed to help the student build the search space (the set of all attribute-value pairs). The learner has to discover how to explore it. By applying a given set of operators and by watching the outcome, he can test (i) which information is relevant, and (ii) to what it is relevant (to syntax or morphology). However, in this kind of dialogue (controlled trial and error) the system not only answers the questions asked by the learner, but also assists him in determining what questions are meaningful in this context.

Learning, be it by man or by machine, implies exchange of information between two systems, for example, a native speaker (expert) and a foreigner (learner). We will start by describing some of the features our system needs to have in order to allow for such an information exchange. We will then give a detailed example, showing what such a dialogue between a human learner and the machine might look like. Finally we will discuss whether machines can acquire linguistic competency in a humanlike way.

Before showing how the system is designed to work, let us specify more clearly what the learning objective is.

4 THE STUDENT'S LEARNING OBJECTIVE:

The learner's task consists of incrementally learning the morpho-syntactic rules of personal pronouns in French. More precisely, the student is expected to acquire the necessary knowledge in order to generate sentences composed of several pronouns (see examples (a) - (i)). In order to achieve this goal, he has to learn:

- how to express a given concept (morphemes),
- how to linearize these concepts (sentence patterns), and
- under what conditions (rules) to use each of these words or sentence forms.

MORPHOLOGY Example of rules to determine MORPHOLOGY

SPEAKER: je, me, moi, - nous	if	SYNT.FUNCTION: direct object
LISTENER: tu, te, toi, - vous		PERSON: third
ELSE: il, elle, ils, elles		REFLEXIVE: no
le, la, les, lui, leur		QUANTITY: definite
on, en, se, soi, eux		NUMBER: singular
		GENDER: female

then DIRECT OBJECT --> la

SYNTAX:

a) S-DO-IO-V	je la lui présente	I introduce her to him
b) S-IO-DO-V	je te la présente	I introduce her to you
c) S-DO-V-PP-IO	je te présente à elle	I introduce you to her
d) S-IO-V-PP-IO	je lui parlerai de toi	I will talk her about you
e) V-DO-IO	présente-la moi	Introduce her to me
f) neg-DO-IO-V-neg	ne la lui présente pas	Don't introduce her to him
g) neg-IO-DO-V-neg	ne me la présente pas	Don't introduce her to me
h) neg-DO-V-neg-PP-IO	ne me présente pas à elle	Don't introduce me to her
i) neg-IO-V-neg-PP-IO	ne lui parle pas de moi	Don't talk her about me

S: subject, DO: direct object, IO: indirect object, pp: preposition, neg: negation, V: verb

As one can see from the data, pronoun-constructions in French can be fairly complex (4). This complexity is due to:

- * the number of features necessary to determine word order or morphology:

PART OF SPEECH: (noun, pronoun)
je parle Pierre (noun)
je lui parle (pronoun)

SYNTACTIC FUNCTION (subject, direct object, indirect object)
il écrit à Pierre (subject)
Paul lui écrit (indirect object)

SENTENCE-TYPE: (declarative, interrogative, command)
tu me le donnes? (interrogative)
donne-le moi! (command)

NEGATION: (yes, no)
donnes-le moi! (positive)
ne me le donnes pas! (negative)

COMMUNICATIVE-ROLES: (I, you, he)
je te LE donne (IO = you)
je LE lui donne (IO = he)

NUMBER: (singular, plural, indefinite)
je te le garde (singular)
je te les garde (plural)
je t'en garde (indefinite)

GENDER: (male, female)
je le vois (male)
je la vois (female)

VERB CONSTRUCTION: (type of complement (DO vs IO), type of preposition, reflexivity)
je vois Marie --> je la vois (direct object)
je parle à Marie --> je lui parle (indirect object)

SEMANTIC FEATURES: (animate, inanimate)
il m'emmène à Rome --> il m' y emmène
il me présente à sa mère --> il me présente à elle

* the structure of these features: if one compares (a) and (c), one will notice that the form of the indirect object (lui vs elle) depends on the value of the direct object (horizontal dependency);

* the interdependence of syntax and morphology: practically all variables, except NUMBER and GENDER are relevant both for syntax and morphology. Furthermore, the position of the direct object pronoun may depend on the value of the indirect object (compare (a) and (b) here above). In other words, changes in morphology often imply changes in syntactic structure.

* the various knowledge sources: the determination of morphology and syntax requires information about the referent (number, gender, animacy), text functions (syntactic status of noun-phrase: noun vs pronoun, topicalisation, person), negation (positive/negative), speech-act, (statement/question/command), verb-construction (type of complement: direct/indirect, type of preposition: à, de), etc.

Given these intricacies it is easy to understand why students so often fail to learn these rules. Modelling their learning is thus a challenging task.

5 HOW CAN THE LEARNER BE INTEGRATED INTO THE PROCESS?

If one accepts this view of learning, then the problem of the student is to find out how to build and how to intelligently reduce the search space. The system will help the student in various ways.

First of all, it will answer certain kinds of questions:

- (a) How does one say x ?
- (b) What would happen if...?
- (c) Can one say x ?
- (d) How should one say x ?
- (e) Why does one say x ?

All these questions occur in some form or another in natural settings. The following examples may illustrate these strategies or testing modes:

(a) Question: How does one say: "je lui pense"

Answer: je pense à elle
je pense à lui
je le pense (5)

(b) Question: What would happen, if in the following sentence: Paul parle à Marie (Paul talks to Mary) the object-noun was pronominalized?

Answer: Paul lui parle

(c) Question: Can one say: "je lui pense"?

Answer: no

(d) Question: Instead of saying "je lui pense", how should one say?

Answer: je pense à elle
je pense à lui
je le pense (5)

(e) Question: Why does one say: "Je le pense"

Answer: explanation given by the system

These strategies are complementary in that they correspond to different learning needs. They provide different kinds of feedback. The first two methods (the inductive approach) seem useful if one does not have much knowledge yet. The third one allows to test the degree of generality or the extension of a given rule (deductive reasoning), the fourth method provides additional information in case of incorrect performance, while the last question may either confirm a hypothesis, or correct a misconception.

Second, the system should show how to reach the solution (the demonstrative mode). This might be helpful if the student gets stranded, not knowing what to do. In this case the system takes over, showing how information may be processed. By watching the system, the student may learn how to explore, i.e., how to generate and test a set of hypotheses.

Third, the system keeps a record of the whole dialogue. Such a trace has many advantages: it allows the student to verify, to explain and to remember. He may thus (i) check the consistency of the rules, (ii) justify a given conclusion in the light of evidence and (iii) reorganize his knowledge base. This last possibility should enhance his perception of underlying regularities.

Psychologists could use this trace to infer the student's learning strategies. The rules a student has been testing at a given moment may be inferred on the basis of the nature and order of the questions being asked.

Finally, teachers could use the trace-function to gain feedback concerning the order of presentation of the data. By varying the nature and order of information, they can determine experimentally the complexity of the data (examples, rules), and thereby the relative efficiency of various teaching-strategies.

6 THE FUNCTIONING OF THE SYSTEM:

The program works interactively. The user is given a set of options from which he has to choose. The system converts this input into the adequate output, i.e., linguistic form. Input are meanings (what to say), output are sentences (how to say it).

The process is started with a list of nouns and verbs. This list is a kind of knowledge base, i.e., a set of facts a potential user may want to talk about. This base is limited in scale, and arbitrary, in that it is given by the system. However, this limitation is easily overcome. The base can be extended by the user at any moment. The important point is that, by feeding nouns and verbs into the knowledge base and by choosing among these entities, the student signals what he wants to say. In doing so, he builds propositions of various complexity (one-, two-, or three place predicates).

The system will operate on these structures and build simple declarative sentences. In other words, at this stage of interaction it is assumed that the student wants to know how the intended meaning translates into this canonical form. For example, the input (a) would yield the output (b).

STUDENT input:(a)	SYSTEM output:(b)
regarder (Manuel, Christine)	==> Manuel regarde Christine
watch (Manuel, Christine)	==> Manuel watches Christine

The student is queried again to determine what he wants to say. Basically he has two possibilities. Either he tries a complete new idea (proposition), or he modifies part of the preceding one. In this latter case, the system provides a list of options (attribute-value pairs), inviting the student to discover what happens, i.e. how morphology and/or syntax are affected, as he changes the value of any of the attributes such as PART OF SPEECH, SENTENCE MODE, NEGATION, and so forth. Let us assume that the student had chosen to replace respectively Manuel and Christine by a pronoun. In this case the system would produce the following sentences:

Il regarde Christine
Manuel la regarde

By comparing these sentences with the base form, the student should notice certain differences and draw the necessary conclusions. For example, given the data he may conclude that:

- R1: if the direct object is pronominalized, then it moves in front of the verb (syntax).
- R2: case (syntactic function) is morphologically relevant: if the subject is pronominalized then its surface form is "il".
- R3: if the direct object is pronominalized then its surface form is "la".

Control is returned to the user. Actually, from now on we are in a loop, with the dialogue having basically the same form. However, in each cycle the hypothesis to be tested is likely to be different and it is interesting to watch how a student proceeds in acquiring competency. What does he want to know? Is he systematic? What kind of strategy does he use (breadth first, depth first etc.)? Under what conditions does he change his method? etc.

The learner's problem is three-fold, he must find out:

- * which parameters (attributes) are relevant,
- * to what linguistic component they are relevant (syntax and/or morphology), and
- * to what extent they are relevant (6).

A student may thus want to know:

- * whether the variable GENDER is morphologically relevant,
- * whether this is the only relevant variable, or if other variables come into play;
- * whether it is relevant for all cases, irrespective of, for example, communicative role, negation or sentence mode (compare (e) and (g));

It should be noted, that every time the student is given control, he can choose two things: (i) the kind of information he wants to convey (what to say), and (ii) the dialogue-mode, i.e., HOW DOES ONE SAY?, CAN ONE SAY?, etc.). The following diagram illustrates the information flow.

insert figure 1 here

This kind of environment has three basic functions:

- (a) to answer different kinds of questions,
- (b) to convert meaning into form, and
- (c) to help the student to discover how changes in meaning are reflected in changes in form.

It should be noted that the student has most of the control. The following examples should give an idea of the dialogue. These hypothetical dialogues

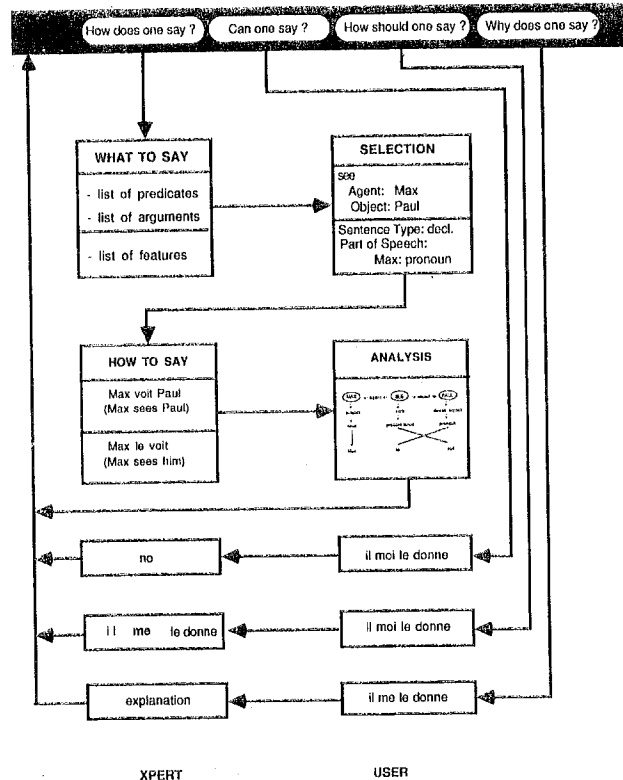


FIGURE 1

serve illustrative purposes. However, we believe that they are reasonably close to what might be encountered in an experimental session.

6.1 EXAMPLE DIALOGUE NUMBER 1:

The student's question (dialogue mode) is: HOW DOES ONE SAY? The figure below contains three columns which express respectively the student's intentions, i.e. what he wants to say, his observations, and his conclusions with respect to syntax and morphology.

insert figure 2 here

Having generated the following proposition:

voir (Max, Paul)
(see (Max, Paul))

he wants to know what would happen, if both arguments (Max, Paul) were pronominalized. The system generates the following answer:

(1) il le voit

The student analyzes this sentence and draws as conclusions Rule 1 and Rule 2, mentioned here above. He goes then on to ask what would happen if PAUL was replaced by MARY. The system answers:

(2) il la voit

The student concludes that GENDER is not relevant with regard to word order, but is a necessary condition to determine morphology (Rule 3). This latter kind of knowledge could be expressed as:

R3: if PART OF SPEECH: pronoun
& SYNTACTIC FUNCTION: direct object
& GENDER: female
then PRONOUN: la
else
if GENDER: male
then PRONOUN: le

In the next question he is concerned with the relevancy of NUMBER. He asks: what would happen if the direct object were CHILDREN (les enfants)? The system's answer

(3) il les voit

allows him to conclude that NUMBER is relevant for morphology but not for syntax, as there are no changes in word order, but there is a change in form. This fact is encoded in the following rule:

(INPUT) HOW DOES ONE SAY	(OUTPUT) OBSERVATION	CONCLUSION
1) voir (Max, Paul) Max = pronoun Paul = pronoun → il <u>la</u> voit	DO precedes V	<u>Syntax</u> SYNTACTIC CATEGORY and the SYNTACTIC FUNCTION of the referent are syntactically relevant R1: if SYNTACTIC CATEGORY: pronoun if SYNTACTIC FUNCTION: dir. obj. then pronoun in front of the verb Subject - Direct Object - Verb
	pronoun = la	<u>Morphology</u> SYNTACTIC FUNCTION is morphologically relevant R2: if SYNTACTIC FUNCTION: subject then PRONOUN: il if SYNTACTIC FUNCTION: dir. obj. then PRONOUN: la
2) voir (Max, <u>Maria</u>) Maria = pronoun → il <u>la</u> voit	position of DO consistent with R1	<u>Syntax</u> GENDER is syntactically not relevant
	change in form	<u>Morphology</u> GENDER is morphologically relevant R3: if SYNTACTIC CATEGORY: pronoun if SYNTACTIC FUNCTION: dir. obj. if GENDER: female then PRONOUN: la if GENDER: male then PRONOUN: le
3) voir (Max, <u>enfants</u>) enfants = pronoun NUMBER = plural → il <u>les</u> voit	no change in position	<u>Syntax</u> NUMBER is syntactically not relevant
	change in form	<u>Morphology</u> R4: if SYNTACTIC FUNCTION: dir. obj. if NUMBER: plural if GENDER: male (*) then PRONOUN: les
4) <u>parler</u> (Max, <u>Marie</u>) Max = pronoun Marie = pronoun → il <u>lui</u> parle	IO precedes V see R1	<u>Syntax</u> Syntactic function (case) is relevant R5: if SYNTACTIC CATEGORY: pronoun if SYNTACTIC FUNCTION: indir.obj. then: Subject - Indir.Object - Verb Generalization of R1 & R5 = R6: R6: if an object is pronominalized then: Subject - Object - Verb
	change in form	<u>Morphology</u> SYNTACTIC FUNCTION of IO morphologically relevant (see R2) R7: if SYNTACTIC FUNCTION: indir.obj. if GENDER: female (**) then PRONOUN: lui
5) parler (Max, <u>Paul</u>) Paul = pronoun → il <u>lui</u> parle	consistent with preceding rules	<u>Syntax</u> GENDER is syntactically not relevant
	no change in morphology	<u>Morphology</u> The gender of the IO is morphologically not relevant, consequently relax the GENDER constraint of R7 Correction of R7: R8: if SYNTACTIC FUNCTION: indir.obj. then PRONOUN: lui

(*) since GENDER was relevant for the singular the student assumes that it is also relevant for the plural

(**) since GENDER was relevant for the IO (R3) the student assumes that it is also relevant for the IO

FIGURE 2

R4: if SYNTACTIC FUNCTION: direct object
& GENDER: male
& NUMBER: plural
then PRONOUN: les

It is interesting to notice, that this rule is too specific, because GENDER is not a necessary condition. However, this conclusion is perfectly reasonable given the data encountered so far. GENDER was a necessary condition for singular (see rule 3), and since then there has been no evidence to the contrary. Consequently, the student has no way to conclude from the data, that for direct objects GENDER is generally relevant only for the singular. (The only reason we could think of that a student might consider this last hypothesis, would come from his knowledge of another language which has the very same property.)

It is also noteworthy that for objects, GENDER is only relevant for the SINGULAR. This has procedural implications; namely that NUMBER should be processed prior to GENDER. The former being more informative than the latter.

In the following cycle (sentence 4) the student changes the proposition altogether, asking the system how one would say:

parler (Max, Paul)

when both arguments are pronominalized. This would yield the following sentence:

(4) Il lui parle

From that he may conclude that the indirect object precedes the verb (Rule 5). Recognizing the similarity with rule 1, i.e., recognizing the fact that the syntactic status of the object (direct vs indirect) does not affect word order, he may generalize these two rules and replace them by rule 6:

R6: if an object is pronominalized, it precedes the verb

This rule is more general than the former ones, in that the distinction between direct and indirect object has been dropped. It should be noted, however, that this rule, even though correct in the light of evidence, i.e., data encountered so far, is too general. For example, it does not apply to sentences composed of two objects (three place predicates). In other words, this rule needs refinement, i.e., additional constraints.

With respect to morphology, the student concludes that the attribute CASE (syntactic function) is relevant, which yields the following rule:

R7: if SYNTACTIC FUNCTION: indirect object
& GENDER: female
then PRONOUN: lui

Again, the morpheme is overspecified, because GENDER is not a necessary condition. Having noticed that GENDER was relevant for direct objects (rule 3) the student has overgeneralized, assuming that it was also relevant for the indirect object. It is noteworthy, however, that this particular overgeneralization does not produce incorrect results.

Finally, the student asks the system to replace MARY by PAUL. Getting the same answer as in 4, he concludes that for indirect objects the GENDER is irrelevant for syntax as well as for morphology. Consequently, he relaxes the gender-constraint of rule 7. Once again, this conclusion is valid only with respect to the set of examples he saw.

6.2 EXAMPLE DIALOGUE 2:

This time the dialogue-mode is CAN ONE SAY. The three columns correspond to the student's questions, his hypotheses, and his conclusions. The controlled variable (a change of attribute or a change of its value) is underlined.

insert figure 3 here

The figure being rather self explanatory, We will make only some short comments. At stage 3 the student wants to know whether the communicative role of the indirect object, the attribute PERSON, is syntactically relevant. From the data he has seen, he concludes that this was not the case. However, this conclusion, even though correct with respect to the data, has to be revised in the light of new evidence (next sentence, i.e., sentence 4).

It is interesting to note, that the student would probably never have drawn this conclusion if sentence 4 had preceded sentence 3. In other words, he would have noticed the relevancy of the attribute PERSON right away, and never have drawn conclusion 5.

il le LUI donne
il TE le donne

This shows how the order of the data is a critical variable determining the efficiency of rule-inference, i.e., what conclusions are drawn at what moment.

7 CAN MACHINES ACQUIRE LINGUISTIC COMPETENCY IN A "HUMAN" WAY ?

Actually there are three questions:

- Can machines learn?
- Can they learn in an intelligent or "human" way?
- What kind of knowledge would a computer program need to have in order to learn the rules I have been talking about?

The answer to the first question is clearly yes (see Michalski, Carbonell & Mitchell 1983). The latter two questions are more controversial. Let us begin with the last one.

Inductive learning basically consists of drawing conclusions from the similarities and differences of abstract data descriptions (contrastive analysis). The crucial points are thus data description and analysis:

- in what terms should we characterize the data?
- what additional kind of knowledge is needed to infer the rules ?

CAN ONE SAY ?	HYPOTHESIS:	CONCLUSION:
1 il me le donne yes	Q1: Do both objects precede the verb ?	Answer to Q1: yes, both objects precede the verb
	Q2: Which one precedes the other ?	Answer to Q2: the indirect object precedes the direct one: S-IO-OO-V
1a il me LA donne yes	Q3: Is the GENDER of the direct object syntactically relevant?	Answer to Q3: the GENDER is syntactically not relevant
2a il me LES donne yes	Q4: Is the NUMBER of the direct object syntactically relevant?	Answer to Q4: the variable NUMBER is not relevant for syntax
3 il LE le donne yes	Q5: Is PERSON of the indirect object syntactically relevant ?	Answer to Q5: no, PERSON is not relevant for syntax
4 il LUI le donne no: il le lui donne		Correction and refinement of Conclusions 2 and 5: the variable PERSON is syntactically relevant: Conclusion 6: if PERSON-IO: 3d then: S-DO-IO-V Conclusion 7: if PERSON-IO: 1st or 2nd then: S-IO-OO-V
5 il le SE garde no: il se le garde		refinement of conclusion 6 Conclusion 8: if PERSON-IO: 3d if V-CONSTRUCT.: reflex. then: S-IO-OO-V
6 il s' EN moque yes		confirms conclusion 8 conclusion 9: if NUMBER-OO: indefinite then: S-IO-OO-V
7 il en LUI donne no: il lui en donne	Q6: If the OO = en & if the IO = lui, Which one of them precedes the other ?	Answer to Q6: the indirect object, consequently conclusion 6 has to be refined. Conclusion 10: if PERSON-IO: 3d if NUMBER-OO: indefinite then: S-IO-OO-V
8 il m' en donne yes		confirms conclusions 7 & 9 with regard to the examples given in 7 and 9 we may relax the PERSON-constraint of conclusion 10

FIGURE 3

Obviously, a system capable of performing the kind of learning we have been talking about would have to be able to parse the sentences; that is, it would have to produce as output an adequate description of the input sentences described above.

This raises a terminological problem. Data can be described in various ways. Different descriptions can be functionally equivalent (7). Clearly, the choice of metalinguistic terminology differs depending on whether the goal is machine learning or modelling "human" learning. In the first case, the problem is descriptive adequacy, whereas in the second case we deal with an additional constraint, that of the universal status of the terminology. Do all humans, irrespective of culture and education, use the kind of terms linguists use to analyze sentences? Is there a universally shared subset of metalinguistic vocabulary? In the absence of answers to these empirical questions we will stick with the terminology currently used in computational linguistics.

A different, but related problem is the question of how a system may be enabled to draw conclusions from a set of data (inferring general rules).

As we have said above, generalizations are made on the basis of contrastive analysis. In order to allow for such generalizations, the learning component needs a hierarchically structured metalanguage, that is, a vocabulary whose low level concepts (primitives) are subsumed by more highly ordered, abstract forms of knowledge. For example:

masculine & feminine ==> GENDER;
singular & plural ==> NUMBER;
subject, direct object ==> CASE

We will now turn to the question of whether computers can learn in an intelligent or "human" way? Obviously this question raises the problem of what intelligence is. Instead of answering this question, we will focus on two aspects of intelligent learning, namely economy and flexibility of methods.

Exhaustive search is neither natural nor economical. Since memory is associative, we find it hard to be consistently systematic. Like gamblers, we tend to use search methods which are more or less risky.

People (learners) generally have a set of methods and a separate component (critique) for evaluating these strategies with respect to their relative efficiency. As different problems require different problem-solving methods, it is very unlikely that there is a unique, universal problem-solving method. People tend to be opportunistic in their approach rather than systematic or scientific. Both the nature of strategies and the depth of processing will vary with the needs of the learner. Corroboratively, it is equally unlikely that one finds the optimal method immediately, since one operates on incomplete data. Inductive learning is typically incremental. Hence methods have to be adapted or gradually refined in the light of new evidence.

Intelligent learning is thus intimately linked to strategic knowledge (8) and to (more or less) general information-processing principles. These principles may be expressed in terms of simplicity, informativeness, generality, and so forth.

For example, the notion of simplicity may be used to choose among different options. In fact, a learner could hypothesize that two-place predicates (to see) are simpler to process than three-place predicates (to give).

The notion of information is related to efficiency. It can be used to reduce the search space. This claim is substantiated by the fact that rules governing morphology of first and second persons (I, you) are generally learned faster than those which determine the form of the third person (he).

In conclusion, we believe that, in principle, certain aspects of intelligent learning could be modelled by a computer. However, before trying to model human learning, it may be worthwhile to start gathering data on how humans learn. This is precisely one of our goals. By watching people as they use this tool, i.e. by keeping a trace of the dialogue, one should be able to infer the strategies they use.

8 CONCLUSION :

We have described a system under development that is meant to be a tool for theory builders (cognitive psychologists), application designers (language teachers) and end users (students). The system is meant to assist psychologists, teachers and students in their respective tasks: model learning, optimize teaching and learning strategies.

The emphasis in this paper has been on learning rather than teaching. For the time being the task of learning is to be performed by a human, however, in principle it is possible to extend the system so as to allow for automatic learning, the ultimate goal being to model human-like behavior.

Computers, with their large, indelible memories, are powerful tools. They allow us to control virtually any number of parameters. Consequently, one can trace a reasoning process or test a given theory, i.e., determine empirically how different variables affect the efficiency of learning, and so forth.

This has an interesting consequence with respect to theoretical commitments. Instead of claiming an all-encompassing model or theory, one can write a program general and flexible enough to permit the testing of various theories. That is what we are trying to do.

Watching how people use the tool, we may gain insights about the way humans learn (strategies), and thus eventually move from artificial to natural intelligence.

NOTES :

(1) This message has to be interpreted. Thus the learning task is not the surface form of the message, i.e., words and sentences, but the underlying principles (abstractions: rules and sentence patterns) allowing their generation. While some forms (e.g., words) have to be learned, they generally serve for illustrative purposes. Rote learning of the entire set of surface forms (words and word combinations) is not only inefficient, but in fact impossible, because of time constraints: there are more possible combinations than we have time to learn.

Learning is thus more than a quantitative change of performance (speed, number of errors). It generally implies a restructuring of the knowledge base.

(2) It should be noted, however, that we are not dealing here with children learning a first language. Instead we would like to model some aspects of the scientific-minded foreign language learner.

(3) One may object that there is a global goal, namely learning the language. However, it seems to me that the primary goal is communication rather than attaining a local objective like, let us say, learning the pronoun system in French.

(4) For a more detailed discussion, in particular with respect to the procedural implications, see Zock, et al. (1986).

(5) In an ambiguous situation the system will either produce all cases (see here above), or ask for clarification. For example:

Student: How does one say: il te me présente
System: This depends on what you want to say.

Do you mean (a) or (b)?
(a) il te présente à moi
(b) il me présente à toi

(6) This last problem, which consists in finding the right degree of generality (underspecification vs overgeneralization), is particularly delicate in that conclusions have to be reached on the basis of incomplete data (incremental learning).

(7) This fact is illustrated by the variety of parsers. Parsers analyze sentences and assign them descriptions on various levels such as: part of speech, syntactic function, case-roles and so forth. For a review of the state of the art see King (1983) or Winograd (1983). For a French parser see Francopoulo (1986).

(8) These strategies could either be part of the system, in which case they must be explicit (one needs a model), or they could be part of the learning process, in which case the system learns not only domain specific knowledge but also methods of how to learn (metaknowledge).

REFERENCES:

- Francopoulo, G.
1986 Machine Learning as a Tool for building a Deterministic Parser in: Rollinger, C. & Horn, W., Second Austrian Congress on Artificial Intelligence, GWA I 86, Springer Verlag
- King, M.
1983 Parsing Natural Languages, Academic Press, New York, 1983
- Michalski, R., Carbonell, J. & Mitchell, T.
1983 Machine Learning: An Artificial Intelligence Approach, Tioga Publishing Company, Palo Alto, California
- Winograd, T.
1983 Language as a Cognitive Process, Addison-Wesley, Reading, Mass.
- Zock, M., G. Sabah, C. Alviset
1986 From Structure to Process: Computer Assisted Teaching of various Strategies for Generating Pronoun-Constructions in French, in: Proceedings of the 11th COLING, Bonn