The function of examples in learning a second language from an instructional text

Carol E. Moon and Steven L. Lytinen
Artificial Intelligence Laboratory
The University of Michigan
Ann Arbor, MI 48109

Abstract

This paper addresses the role that examples play in instructional learning. We discuss several roles that examples can serve when they complement an instruction. We provide functional evidence for some of these roles, arguing why instructions and examples are both necessary for efficient learning. We present a system that learns from instructions which are enhanced by examples. The system, ANT (Acquisition using Native-language Transfer), learns a second language by reading instructions about grammatical rules of the second language as well as examples which use these rules. Finally, we argue for the functional utility of examples in instructional learning on more general grounds, showing how such a strategy can be applicable to other domains besides second language learning.

1 Introduction

People learn a great deal of information by being told it. Furthermore, they can readily assimilate this information with what they already know, allowing them to use it to improve their performance in tasks such as problem-solving or planning. For example, if you are learning calculus, and I teach you about a new integration technique, such as integration by parts, then if all goes well you will be able to use that rule to help in your performance of solving integrals.

However, learning from instructions is not quite that simple. If we look at most instructional textbooks, we find that often instructions about a new rule are accompanied by examples of how that rule can be used. A typical calculus book presents integration by parts by first stating the rule in a succinct manner, and then presenting several example problems, including a step-by-step description of how to solve them using integration by parts.

This paper will address the role of examples in instructions. Why are they necessary? At first glance, they seem superfluous. Don't the instructions provide the learner with all the information that is needed? For integration by parts, doesn't the formula alone say it all?

We will argue that in order to apply new information to some task, often the learner needs other information, which is not easily conveyed through instructions. In particular, the learner needs to know how this new information connects up with what he already knows. This is just one role that examples can play in instructions: they can provide the learner with an "experience" that will help to delineate the role of the new information relative to information which is already known.

We are building a program which learns from instructions augmented with examples. The program, called ANT (Acquisition using Native-language Transfer), operates in the domain of second language learning. This domain is a good one for studying the role of examples in instructions. Second language learners seem to "transfer" their knowledge of their native language over to the second language, modifying rules according to the differences between languages (Gass, 1980; Lado, 1957; Selinker, 1983). As a result, incorporating new rules into existing linguistic knowledge is a major part of second language learning. As we will see, our program uses the examples that accompany instructions to determine which English rule(s) need to be modified for German.

ANT begins with a knowledge base of grammar rules and lexical entries for English, which enable it to understand English descriptions of grammar rules. The program also develops a knowledge base of rules for German. Initially, the program assumes that all of its knowledge for English will apply to German. In this way, it expects to be told about differences between the two languages so that it can modify its German rules accordingly. ANT receives input which we have taken from an introductory German textbook. The input consists of instructions about German grammar rules, as well as examples of German sentences that illustrate these rules. Here is a typical input to the system:

> In German, verbs come at the end of relative clauses.
> Example: Der Ameisenbär, der die Amiese fraß, ißt Kaviar auch.
> (The aardvark who ate the ants eats caviar, too.)

ANT modifies its knowledge base of grammar rules for German according to the instructions and examples, and can then understand German sentences which it has not seen previously that use the new construction. In this paper, we will show how ANT utilizes the information in the examples to clarify the meaning of the instructions.

2 The roles of instructions and examples

Examples can function in several distinct ways when they are used to complement instructions. In this section we will discuss several of these roles, and the psychological evidence that exists for many of them. In the next section, we will show how examples serve analagous roles in our computer model.

2.1 Examples

Retrieval cues

Examples can provide additional retrieval cues for the information in an instruction (Reder,Charney, and Morgan, 1986). For example, a learner may forget the original instruction when he is trying to formulate a relative clause. But the learner may be able to remember the example and hence the lesson of the instruction by remembering the example about the aardvark and caviar.

Overcoming abstraction

Often, a learner needs an example which is less abstract than the instruction. Abstraction can be difficult especially when the learner is unfamiliar or inexperienced with the domain (Reder, *et al.*, 1986). For example, if a student is not particularly knowledgeable about grammar rules explicitly,

a rule like the relative clause one above will perhaps be too abstract for the student to remember. In other words, his lack of command of grammar rules may inhibit the incorporation of this knowledge into his knowledge about using language. When such a student is given an example which serves as an instance of the application of the rule, he may more easily see both what is meant by a relative clause and which verb is moved. The example can invoke his rules about clauses and illustrate how his original notion of a relative clause must be changed.

The student probably knows many examples of relative clauses without necessarily having a command of the grammatical terminology. In this way, examples avoid the abstraction of the instruction, allowing the learner to build his own generalization after several examples.

Reasoning by analogy

Related to the situation of the student above is the idea that when an instruction is too abstract, examples may allow a learner to solve the next problem (i.e., understand or formulate a subsequent sentence) by analogy. Analogy has been shown to be an important problem solving and skill acquisition method, especially when the learner has little domain knowledge, like in the situation above where the hypothetical student's lack of explicit grammar knowledge (Pirolli and Anderson, 1985).

Modifying prior knowledge

Examples can also be used as a way to discover what existing knowledge must be modified according to the instruction. How is the instruction related to previous knowledge about the domain? In which contexts should the new instruction be used? Examples can show the system how the instruction is applied and in which contexts it should apply (Reder *et al.*, 1986; Stein and Bransford, 1979). Instructions alone usually do not provide this information. How examples perform this role in ANT will be shown in the next section.

Filling in details

Finally, examples can instantiate details not made explicit in the instruction (Reder *et al.*, 1986). Suppose the system was given the following instruction:

> In German, to express 'to like' the verb 'haben' is used with the adverb 'gern.'

Much information that the system would need to build an appropriate rule for this construction would be missing. Where does the object of 'like' (or 'haben') go? What is the relative ordering of 'haben' and 'gern'? Subsequent examples would allow the system to deduce that the noun phrase that is the object comes between 'haben' and 'gern' and that 'haben' precedes 'gern.' ANT utilizes this function of examples to a great extent. Introductory language textbooks seem to rely on this as well, providing the student with rather brief instructions that are supplemented by a series of examples, from which the student must learn the complete details of such a construction.

2.2 Why do we need instructions?

Given the roles of examples discussed above, one might be led to believe that instructions play little or no role in the learning process. Perhaps the learner completely ignores instructions, paying attention only to the examples which accompany them. However, even given the above roles for examples, instructions still can have important roles to play:

Focus

Instructions give the learner clues about what features of the examples are important. If the learner were given no instruction, how would he decide which features of the example to focus on? In our relative clause example, he could just as easily notice the agreement in case and gender between the noun and the determiner as a feature to be learned as he could the different word order in the clause. The instruction thus can make the significant processing of the example more efficient.

Expectations

In ANT, examples play a role in changing the system's expectation of a relative clause. Because ANT assumes that any rules it knows about English apply to German unless it learns otherwise, it tries to use its grammatical and semantic rules to parse the German examples. But certainly in the relative clause example, the German example will not be successfully parsed using the English rules because the constraints on word order in the grammar would not match the constraints determined by the word order in the example. ANT uses the instruction to alter its expectations. It determines that the instruction is focusing on word order in relative clauses. When ANT subsequently parses the German examples using relative clauses, it relaxes word order constraints, thus allowing the parser to build a parse tree and representation of the example. This process will be explained in more detail later.

3 An example of ANT's learning

We have seen many possible roles for examples and instructions to play in the learning process. Let us now turn to ANT, and see how these roles come into play in our computer model. We will see that the roles are essential to the learning process, providing good functional explanations that complement the psychological evidence discussed in the previous section.

All linguistic knowledge in ANT is represented using a unification-style grammar (Shieber, 1986)[1]. In this approach, word order information, as well as information about the functional relations between words, is explicitly and declaratively represented. This is important for the system's task, because it must be able to manipulate various components of English rules in order to form new German rules. Another key feature of this approach is that the structure of grammar rules used by ANT is the same as the structure ANT produces in parsing. (for details, see Lytinen and Moon, 1988). As we will see, this allows ANT to extract new rules from its understanding of examples that it is presented.

Let us now consider the entire process which takes place when ANT learns a new rule. The example we will discuss is our relative clause example:

> In German, verbs come at the end of relative clauses.
> Example: Der Ameisenbär, der die Ameise fraß, ißt Kaviar auch.
> (The aardvark who ate the ants eats caviar, too.)

The representation which ANT produces when it reads the instruction is shown below:

[1] Part of the information in unification rules is analagous to the phrase structure information which is often encoded in context-free grammars. For the sake of simplicity, we will use the context-free notation in this paper, even though ANT uses the unification-style versions of these rules.

Order:
Location : RelClause (**RELC**)
Constituent : Verb (**V**)
Position : last

As ANT processes an instruction, it determines what kind of instruction it is. Does the instruction refer to word order, or does it refer to grammatical features, like case, gender, or number? For the relative clause instruction, ANT categorizes it as a **REORDER** instruction, for it is an instruction about changing word order. After producing the representation above, ANT needs to use what it knows about English relative clauses in order to incorporate this new information. This is because the instruction does not completely describe German relative clauses; it just describes the difference between German and English relative clauses. ANT must retrieve its English relative clause rules, as well as rules about subconstituents of relative clauses, and modify some (or all) of them. As we will see, this is a rather difficult task.

The main problem is that the instruction alone does not tell us which rules to modify. It can cue the system to find rules labeled **RELC** (i.e., rules about relative clauses). But if the surface grammar (embedded within unification rules) of our relative clause rules in English are something like **RELC** → **RELPRON VP**, the relative clause rules are not the ones which need modification. Rather, it is the rules about verb phrases (as they occur in relative clauses) which must be modified, since they are the ones which ultimately specify where the verb will occur in the relative clause. If we simply tried to incorporate the new information into our English clause rule, we would arrive at an incorrect result. Our surface grammar might end up with a rule like this: **RELC** → **RELPRON VP V**. This rule would mean that German relative clauses had two verbs, one inside the **VP**, and the other one at the end of the clause.

One possible strategy for finding out which rule should be modified is to expand all of the clause rule's subconstituents, searching the grammar for rules which refer to a **V**. However, this approach could lead to a very extensive search, since in general the constituent which we are trying to find could be nested arbitrarily deep in the grammar. In the worst case, the system would end up inspecting its entire grammar, searching for the constituent in question. In addition, there is no guarantee that the system will find the correct constituent. A verb can be derived from many different places in the grammar. How can we insure that the correct occurrence will be the one that is found in the search?

This is where the processing of an example comes into play. Instead of performing this search, ANT parses the German example, using its English grammar rules. As we will see, the German example forces the system to use the English grammar rule which must be modified. In particular, the clause and verb phrase rules that need to be altered will be used. In this way, the example brings the rules to be altered to the attention of the learning mechanism. As a result, the potentially exhaustive search for relevant rules is avoided.

In order to process the German example, ANT must relax some of the constraints in its English grammar rules. Otherwise, the parse would fail, because the German example does not conform to the grammar of English. In particular, ordering constraints on constituents are relaxed. ANT knows to do this because it has classified the input instruction as a **REORDER**ing rule. Since

this instruction is about relative clauses, all constituent-ordering constraints are dropped for subconstituents of the category **RELC**. The relevant rules are the following[2]:

RELC → RELPRON VP (1)
VP → V NP (2)

These rules encode word order information in English that a relative clause consists of a relative pronoun followed by a verb phrase (i.e., a verb and a noun phrase).

When the system tries to apply these rules, the order of the right-hand side constituents is instantiated by the word order of the input examples. Relaxing constraints allows ordering information for constituents to be derived from the examples, not from higher rules in the grammar. In general, the instruction focuses the system on a feature like case, word order, or word choice. Then constraints from the English rules are relaxed so that the information from the example can take precedence and determine the actual correct values for those features in the German rules.

Once the example is parsed, the correct rule for German relative clauses is embedded within the final structure. As we mentioned earlier, this is because the structure of grammar rules used by ANT is the same as the structure ANT produces in parsing. ANT extracts the rule from this structure, once again using the fact that the rule it is learning is a **REORDER**ing rule to extract the constituent ordering information from within the relative clause to replace those constraints in the original English rules. It turns out that the order requirements which change are within rule 2 above. Namely, the order of the constituents **V** and **NP** must be reversed. But ANT cannot simply rewrite the rule this way, since rule 2 is not just used within relative clauses. Modifications should only be local to relative clauses, so ANT generates a new category, called **VP1**, as a subconstituent of German relative clauses, in which the **V** is the final constituent. The resulting rules would be the following:

RELC → RELPRON VP1 (3)
VP1 → NP V (4)

4 Conclusion

We have discussed many possible roles that examples and instructions can play in learning. Several of these roles turn out to be essential to the learning process used in ANT. First, we have seen why instructions alone do not provide ANT with enough information to learn a new grammar rule. Without the examples, the system would not know which of its existing rules need to be altered. In other words, ANT cannot readily access its previous relevant knowledge unless it is given examples. In addition, the instruction may also omit details (like the exact ordering of constituents) which are necessary to learn the new rule. Without examples, the rule-inferring process could be prohibitively expensive.

We have also seen that the learning process would be much less efficient without the information provided in the instructions. Although given enough examples, it may very well be possible to

[2]These are the relevant rules for the example we are considering. Other relative clause rules would be modified by other examples.

induce the correct grammar rules for a language (see Berwick, 1985), the process would be much slower without the instructions to guide it. They serve as a high-level guide to focus the ANT's attention on the correct item or feature to be learned.

In the relative clause example, the fact that our instruction states that the rule is about relative clauses allows the system to immediately generalize the German example to all **VP**s which appear in relative clauses. Without this knowledge, the system would not be able to determine from a single example just how general the new **VP** rule should be.

The information provided in the instruction can be thought of as analagous to domain knowledge in explanation-based learning methods (DeJong and Mooney, 1986; Mitchell, Keller, and Kedar-Cabelli, 1986). Domain knowledge allows these systems to generalize to the appropriate level from only one example. Without the domain knowledge, these systems would have to use similarity-based learning techniques, examining several examples before reaching the same generalization. In much the same way, without instructions, ANT would not be able to determine how general its new grammar rule should be without looking for similarities over many examples.

Instructions provide ANT with other essential information. By using the instruction to focus on particular features of the example, ANT can avoid considering all features in the example as being the possible topic to be learned. It can more efficiently process the significant features of the example. ANT also uses expectations derived from the instructions to parse examples by relaxing some constraints in its parsing rules. The relaxed constraints allow information from the example, like word order, to determine the ordering of constituents in the grammar instead of having the grammar dictate the correct ordering.

Our proposed roles for instructions and examples in second language learning should apply to many other learning tasks. In fact, the interaction between instructions and examples that we have outlined ought to be very similar in any learning task in which existing knowledge must be modified for the task being learned. In such tasks, determining which existing rule(s) are affected by the newly presented knowledge could result in a very large search of the existing rule base, unless an example is provided which leads the learner to the affected rule(s). Likewise, similar problems of knowing what features of the example to focus on, inferring details, etc., would be encountered by the learner. This seems to include a very broad range of learning tasks. Examples include the learning of a new card game, in which rules might be expressed as modifications of rules from a card game which the learner already knows; or perhaps learning a new piece of software, such as an editor, in which the learner might rely on knowledge of other similar software that he already knows about.

The interplay between instructions and examples surely is more complex than we have described here. Each type of input could potentially play many other roles in the learning process. These roles may vary, depending on the sort of knowledge being learned, as well as the content of the instructions and examples. The various possiblities of interaction between instructions and examples remain a topic for extensive further research.

References

Berwick, R. (1985). *The Acquisition of Syntactic Knowledge.* MIT Press, Cambridge, MA.

DeJong, G., and Mooney, R. (1986). Explanation-based learning: An alternative view." *Machine Learning 1*, Kluwer Academic Publishers, 1986, Boston, MA, pp. 145-176.

Gass, S. (1980) An investigation of syntactic transfer in adult second language learners." in Scarcella, R., and Krashen, S. (eds.), *Research in Second Language Acquisition*, Newbury House Publishing, Rowley, MA.

Lado, R. (1957). *Linguistics Across Cultures.* University of Michigan Press, Ann Arbor, MI.

Lytinen, S., and Moon, C. (1988). Learning a second language. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, St. Paul, MN, pp. 222-227.

Mitchell, T., Keller R., Kedar-Cabelli, S. (1986) Explanation- based generalization: A unifying view. *Machine Learning 1*, Kluwer Academic Publishers, Boston, MA, pp. 47-80.

Mostow, J. (1983). Operationalizing advice: A problem-solving model. In *Proceedings of the International Machine Learning Workshop*, University of Illinois, June 1983.

Pirolli, P., Anderson, J. (1985) The role of learning from examples in the acquisition of programming skills. in *Canadian Journal of Psychology 39*(2), pp. 240-272.

Reder, L., Charney, D., Morgan, K. (1986) The role of elaborations in learning a skill from an instructional text. in *Memory and Cognition 14*(1), pp. 64-78.

Selinker, L. (1983). Language transfer. in Selinker, L., and Gass, S. (eds.), *Language Transfer and Language Learning*, Newbury House Publishing, Rowley, MA.

Shieber, S. (1986). *An Introduction to Unification-based Styles of Grammar.* CSLI, Palo Alto, CA.

Stein, B., Bransford, J. (1979) Constraints on effective elaboration: Effects of precision and subject generation. in *Journal of Verbal Learning and Verbal Behavior 18*, pp. 769-777.