The Giant: An Agent-based Approach to Knowledge Construction & Sharing

Thomas R. Reichherzer, Alberto J. Cañas, Kenneth M. Ford, Patrick J. Hayes

Institute for Human & Machine Cognition
The University of West Florida
11000 University Parkway
Pensacola, FL 32514
treichhe@ai.uwf.edu

Abstract

The purpose of this paper is to report on a research project of a software agent, called the Giant, to support knowledge construction and sharing among learners. The Giant is embedded in an educational software tool used by students when collaborating across classrooms, throughout schools and countries, on a subject of study. Based on propositions being shared among the students, it automatically draws tentative conclusions and displays them to the user. At the user's request the Giant presents explanations of its reasoning. It is anticipated that through the interaction with this software agent students will review and refine their own knowledge.

Introduction

Over the past three years, The University of West Florida and IBM Latin America have established a communications network of K-12 schools, which includes not only the telecommunications technology, but most importantly a pedagogical methodology, educational software tools, and curriculum material. This network enables students throughout schools in Latin America to collaborate in the construction and sharing of models of their beliefs [Cañas et al. 1996]. The partnership, called Project Quorum, has provided specific tools that enable students to perform this collaboration. One of them, the Concept Map Editor (CMap), is based on the idea of using concept maps as a visual representation of concepts and their relationships as known by an individual [Novak & Gowin 1984]. Concept maps follow the idea of assimilation theory [Ausubel et al. 1978] where meaningful learning, as opposed to rote learning, results from linking new information with relevant preexisting concepts or propositions in the learner's cognitive structure. CMap's graphical interface offers simple point-and-click commands to construct concept maps. The system automatically derives propositions encoded in the concept maps. Optionally, the user can extract propositions manually. All propositions are local to the student's workspace and cannot be seen by others. However, the system enables the user to publish propositions in what is called a 'knowledge soup'—a collection of propositions made by all students. The process of publishing propositions makes claims of an individual visible to others, providing a basis for collaboration in the classroom. Students can see claims from their classmates that are related to their own propositions. (The system has some built-in heuristics to decide about relatedness among propositions.) Furthermore, they can attach messages to propositions made by others, thereby intensifying collaboration in the classroom. At no time does the system reveal any information about ownership of propositions. We believe that this aspect of anonymity will encourage people who are less confident to share their ideas (see [Cañas et al. 1995] for a description of the knowledge sharing tools).

The environment is suitable for a software agent that interacts individually with the student, generating tentative conclusions about topics of study using the students' propositions and a simple set of rules. These claims often represent different threads of thought than those followed by the student. We refer to this agent as the Giant because it 'knows' a great deal, but its conclusions may be unrealistic and its questions therefore silly but amusing. The Giant displays its own claims to the user and lets him/her decide on their correctness. Allowing the student to make these decisions puts the learner into the position of a teacher to the Giant, forcing him/her to review both personal propositions, propositions from others, and the Giant's claims. The system is not guaranteed to draw rational conclusions from the concept map (similar to what a human being would do) or verifying the student's propositions. Such an intention is not pursued in our study.

Preprocessing

The Giant obtains propositions from both the student and the 'knowledge soup.' Propositions are simple sentences typically constructed from two concepts and a link between them. An analysis of the constituents of the sentences is essential for the reasoning process of the Giant. For that purpose we provided the system with a grammar (see Figure 1 below) that closely describes the structure of the sentences derived from the students' concept maps. Our grammar is defined for the Spanish language (most countries in Latin America speak Spanish) with respect to the structure of sentences found in concept maps.

S	•	NP VP
NP	•	[Article I Quantifier I Cardinal] Noun [Adjective]
NP	•	[Article I Quantifier I Cardinal] [Adjective] Noun
VP	•	[Quantifier I Cardinal] VB DO [IDO]
VB	•	[Negation] Verb [Adverb]
VB	•	[Negation] [Adverb] Verb
VB	•	[Modal I Auxiliary] Verb [Adverb]
VB	•	Negation [Adverb] [Auxiliary I Modal] Verb
DO	•	NP
IDO	•	Preposition NP

Figure1: Grammar used by the Giant when parsing propositions.

The grammar also addresses the fact that propositions in a concept map may not always follow the grammar of a natural language. In Figure 2 both propositions "plants some are green" and "plants for example trees" do not build correct English sentences. However, concept maps do not impose any restrictions on how to encode propositions. Therefore, examples, similar to the one in Figure 2, cannot be avoided in a classroom setting.
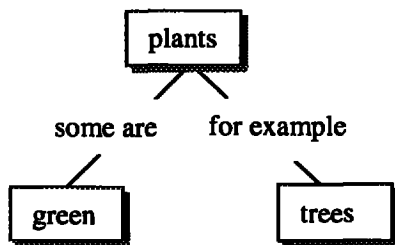

Figure 2: Propositions in a concept map.

The proposed grammar does not match every possible proposition which can be written as a concept map. Our objective is to acquire necessary information from a proposition for the reasoning process of the Giant and not to correctly parse sentences of a natural language. In our grammar, the non-terminal symbols *Article*, *Quantifier*, *Cardinal*, *Noun*, *Adjective*, *Adverb*, *Verb*, *Auxiliary*, *Modal*,

Preposition, and *Noun* denote the corresponding classes of words. The symbol *Negation* represents the Spanish word for negating a proposition.

Using the grammar, we constructed a parser that finds the structure of a proposition and identifies the constituents. The parser takes a proposition as a triplet T=(X, Y, Z) with X as the left concept, Y as the link and Z as the right concept or the right concept followed by another link-concept sequence. The latter case occurs when the proposition extends over more than two concepts and a link (see Figure 3). Each part of the triplet is parsed individually using transition networks [Allen 1987].
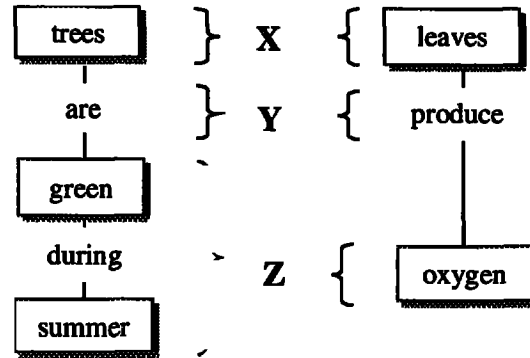

Figure 3: Triplet structure of propositions

The parser uses a lexicon to perform stemming and tagging of the words in the proposition. The result of the parsing process is a table that contains everything essential for the Giant's reasoning. The table reveals information about keywords such as quantifiers, adjectives, adverbs, and verbs found in the link between two concepts. Figure 4 shows an example of a table produced by the parsing process from the sentence "all steam boat captains have moustaches."

nouns in the left concept	steam boat captains
quantifier, cardinal applied to the left concept	all
adjective in the left concept	---
modal/auxiliary verb	---
linking verb	*have*
linking verb (infinitive form)	*have*
adverb	---
right concept	*moustaches*
proposition negated	*no*
number of the left concept	*plural*
voice of the proposition	*active*
part of speech of the right concept	*noun*

Figure 4: Result of a parsing process.

The Giant's behavior and appearance

As a motivational device for children, we have given the Giant a 'personality' of a friendly and eager learner who tries to capture the student's attention by displaying its own 'understanding' of the world. This understanding is based on ideas from the student and the shared knowledge in the classroom. The limited knowledge and the simple reasoning methods utilized by the Giant often lead to less intelligent, sometimes conceptually wrong conclusions. Hence, the Giant appears as a silly, funny, and overall helpless being which sometimes 'knows' some surprising things about the world.

Students may provide clarification to the Giant by telling it with which conclusions he or she agrees or disagrees. Conclusions which are not reasonable at all may be refuted by sending the message 'silly' to the Giant. In some cases the Giant requests new concepts from the user in order to complete propositions (Figure 5 depicts such a request).
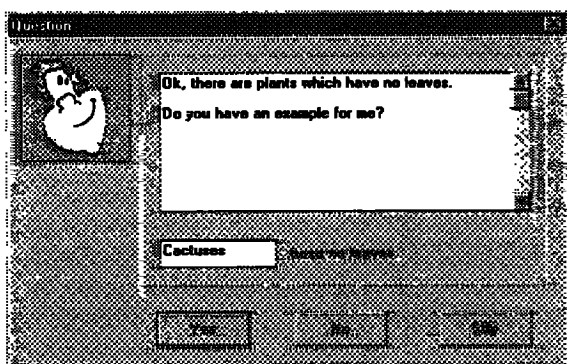

Figure 5: Questions from the Giant. The entry field stating "Cactuses" was completed by the student.

Students may control the Giant's behavior by selecting from the two different settings in the system—the Giant's activity and curiosity. Among these settings, students may choose their Giant to be active or lazy, curious or cautious.

An active Giant is eager to produce as many propositions as possible whereas an inactive Giant provides a small number of propositions that reflect its less active status. The Giant appears curious about the student's work when it applies the entire set of rules to the propositions, yet does not insist on complete fulfillment of the premises of the rules. This behavior setting allows the Giant to derive conclusions that are not necessarily implied in the student's propositions. As a result the Giant might explore new facts about the subject of study. However, in this state there is a good chance that the Giant derives silly statements. The cautious state tries to avoid this by executing rules only if they completely fulfill the premises.

The Giant's curiosity and excitement of learning from the student, and its appreciation when it receives confirma-tion on its own propositions is underlined by facial expres-sions in the form of comical cartoon faces.

The Giant has a general idea of how certain words affect concepts or their relationships. These words, called key-words, are used in the Giant's reasoning process. Since the Giant does not have any common sense knowledge or se-mantic understanding, it cannot draw the line between in-teresting or boring propositions. Instead, the Giant ran-domly selects propositions from the student or the 'knowl-edge soup.' This may cause it to sometimes appear ignorant or moody and to neglect propositions that might be appeal-ing to question.

The Giant never intrudes upon the student. It works in the background producing its own propositions and waits for the student to call it. When asked by the student, it pres-ents an explanation of its reasoning. This is when we enter the interaction between student and Giant. The student has the choice either to supply information to the Giant or to send the Giant into the background. The Giant welcomes the student's advice and uses it to 'clarify' its knowledge about the subject of study. Such behavior is in contrast to the role of an authority that judges and corrects the stu-dent's work.

The Giant's reasoning

For reasoning, the Giant uses a set of rules and the parsed propositions to produce conclusions. The effect of the rules is to generate propositions that are at a tangent to the stu-dent's thread of thoughts and therefore may help the stu-dent to enrich the concept map by adding new concepts or links to it.

We can split the rules into three categories. The first category applies transitivity to propositions coming either from the student or the 'knowledge soup.' The next cate-gory explores keywords that characterize the left concept or the action in the proposition. Such keywords are quantifi-ers, cardinals or adjectives associated with the left concept, adverbs associated with the verb in the link, or verbs that express causal dependencies among the concepts in the concept map. The last category of rules explore the struc-ture of the student's concept maps which may reveal infor-mation about commonalities among concepts.

We use the following notation to refer to the information in the table that the parser creates from a proposition p:

leftCon(p)	the left concept
rightCon(p)	the right concept
linkVerbInf(p)	the linking verb (infinitive form)
linkVerb(p)	the linking verb
voice(p)	the voice of the proposition
modAuxVerb(p)	the modal / auxiliary verbs
quanCardAdj(p)	a quantifier, cardinal, or adjective

	associated with the left concept
adverb(p)	adverb
negated(p)	indicates if p contains the negating word 'no'
speechRCon(p)	part of speech of the right concept

Transitivity

Suppose p_1 and p_2 are two propositions either from the student or the 'knowledge soup.' The two transitivity rules are:

if rightCon(p_1) = leftCon(p_2) **and**
linkVerbInf(p_1) = linkVerbInf(p_2) **and**
voice(p_1) = voice(p_2) **then**
newProposition = leftCon(p_1) linkVerb(p_2) rightCon(p_2)

if rightCon(p_1) = leftCon(p_2) **and**
linkVerbInf(p_1) • linkVerbInf(p_2) **and**
voice(p_1) = voice(p_2) **and**
[(linkVerbInf(p_1), linkVerbInf(p_2)) ∈ Keyword-
list or status = curious] **then**
newProposition = leftCon(p1) linkVerb(p_2) rightCon(p_2)

In both rules, the right concept of p_1 has to match the left concept of p_2 and the voices have to be identical. The difference in the voice matters because in a passive voice proposition the active component is the right concept while in an active voice proposition the active component is the left concept. Hence, we can only apply transitivity to propositions with identical voice.

The first rule matches all propositions p_1 and p_2 with the same root form of the linking verb. The second rule matches propositions with different root forms of the linking verb only if the linking verbs in their root form are in a keyword list or if the Giant is in a curious status. For example, given the two propositions "fish are animals" and "animals breathe air" the Giant concludes that "fish breathe air."

Keywords

The Giant has access to a list of words, or word pairs known to the system as key words that signal the plausibility of a conclusion. The list is fixed; i.e. neither the Giant nor the user can delete or add information in the list. Later systems will explore the potentiality of adapting the list as the Giant's reasoning proceeds.

The construction of the new sentence is influenced by certain words that may occur in p. (For simplicity, we omit details of sentence construction since they are not relevant in the current study.)

Given a proposition p, the following three rules match if a keyword is found or if the Giant is in a curious status.

if [*quanCardAdj(p)*∈ Kewordlist **then**
or status = curious]
newProposition = there are *leftCon(p)* that *linkVerb(p)*
[no(t)] *rightCon(p)*

if [*adverb(p)* ∈ Kewordlist or status = curious] **then**
newProposition = *leftCon(p) linkVerb(p)* not always
rightCon(p)

if [*linkVerbInf(p)* ∈ Kewordlist or status = curi- **then**
ous]
newProposition = the ability to *linkVerbInf(p)* right-
Con(p) requires something

A cautious Giant uses the rules only if a conclusion is plausible due to the existence of a key word in the proposition. For example, from "some plants have leaves" the Giant may conclude that "there are plants that have no leaves," or "the ability to have leaves requires something," or "plants have not always leaves." The latter inference is an example of the Giant's limited capabilities of sentence understanding and construction.

Classification

The classification rule uses the structure of a concept map to perform reasoning. It asserts the presence of a classifier for a set of different concepts when these concepts themselves are linked with identical linking verb to a common concept. The identical linking verbs provide a hint that an enumeration of facts may have occurred. Figure 6 shows an example of a concept map to which the rule applies.
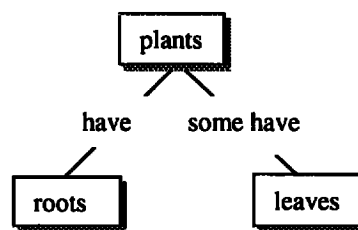

Figure 6: Search for classifier concepts

This concept map enumerates some parts of plants. A possible classifier for the concepts "roots" and "leaves" is "eatable." Thus, the concept map may be extended with "roots and leaves are eatable." The classification rule concludes that "there is a common identifier for roots and leaves." Conditions in the rule for any given propositions p_1 and p_2 are a common left concept, identical linking verbs, and identical voices of the two propositions.

Results

The tables below show the performance of the system on two small school projects with concept maps built in classroom settings by sixth-grade students. In particular, we were interested in what kind of propositions and questions the Giant produces and how students interact with the Giant. The 'knowledge soup' contained propositions derived from all the concept maps. The domains for the two selected projects are "the nature of plants" and "automobiles." Table 1 and 2 present a subset of the student's and the Giant's propositions and consequent questions or explanations by the Giant.

Student's propositions	Giant's propositions
plants need sunlight	there are plants that don't
plants need minerals	have leaves
some plants have leaves	plants make photosyntheses
plants need water	plants release water
	plants don't always have leaves

Giant's questions	Giant's explanations
Do you know which plants do not have leaves? Do you know when plants don't have leaves?	"plants have leaves and leaves make photosyntheses; thus, plants make photosynthesis"

Table 1: Test results of project "the nature of plants."

Student's propositions	Giant's propositions
automobiles are vehicles	there are automobiles that
automobiles need oil	have no wheels
automobiles need gasoline	there are automobiles that
automobiles have windows	do not produce fumes
automobiles have wheels	automobiles produce noise
automobiles produce fumes	

Giant's questions	Giant's explanations
Do you know when automobiles do not produce fumes? Do you know automobiles that have not always wheels?	"automobiles have engines and engines produce noise; thus, automobiles produce noise"

Table 2: Test results of project "automobiles."

In both tests, the student was collaborating with the Giant actively, trying to answer all the Giant's questions. On average, the student refuted two out of ten questions with "silly" and supplied for three out of six questions new concepts.

Summary

We have developed a software agent for Project Quorum, called the Giant with working versions in English, Spanish, and Portuguese. The Giant's task is to draw conclusions from the student's propositions and knowledge that is shared in the classroom, display them to the student, and ask for verification. Our tests have shown that the Giant produces claims that mostly are interesting and smart but sometimes are boring and silly. Most important, however, is that the Giant's claims can suggest new ways of thinking which arise from other propositions. As a consequence, students are faced to analyze their own or other peoples' propositions, leading them to refine and broaden their own knowledge. We anticipate that this process will result in richer and broader concept maps.

We have introduced a relatively small set of inference rules with surprising results. The effort remains to enlarge significantly the set of rules that perform the automatic reasoning and to incorporate learning into the system that would enable the agent to test new facts against those discovered in previous sessions.

References

Allen, J. (1987). *Natural Language Processing*. New York: The Benjamin/Cummings Publishing Company.

Ausubel, D. P., Novak, J. D. Hanesian, H. (1978). *Educational Psychology: A Cognitive View* (2nd ed.). New York: Holt, Rinehart & Winston. Reprinted, 1986. New York: Warbel & Peck.

Cañas, A. J., Ford, K., Hayes, P. J., Brennan, J., Reichherzer, T. (1995). Knowledge Construction and Sharing in Quorum, *Proceedings of AI-ED 95*, Washington DC, pp. 218-225.

Cañas, A. J., Ford, K., Hayes, P. J., Brennan, J., Reichherzer, T., Suri, N. (1996). An Environment for the Construction and Sharing of Knowledge, *Proceedings of the Florida Artificial Intelligence Research Symposium (FLAIRS 96)*, Key West, pp. 238-242.

Novak, J. D. & Gowin, D. B. (1984). *Learning How to Learn*. New York: Cambridge University Press.