TAILORING IMPORTANCE EVALUATION TO READER'S GOALS: A CONTRIBUTION TO DESCRIPTIVE TEXT SUMMARIZATION

Danilo Fum(*), Giovanni Guida(+), Carlo Tasso(˜)

Istituto di Matematica, Informatica e Sistemistica
Universita' di Udine, Italy

ABSTRACT

The paper deals with a new approach to importance evaluation of descriptive texts developed in the framework of SUSY, an experimental system in the domain of text summarization. The problem of taking into account the reader's goals in evaluating importance of different parts of a text is first analyzed. A solution to the design of a goal interpreter capable of computing a quantitative measure of the relevance degree of a piece of text according to a given goal is then proposed, and an example of goal interpreter operation is provided.

INTRODUCTION

Importance evaluation is one of the major issues in text understanding. Human readers, in fact, rank each new piece of information obtained from a text in a sort of importance hierarchy. The mental representation of the meaning of a text can not therefore be assumed to be flat, objective, and reader-independent, but it generally contains a lot of subjective judgmental knowledge. Importance evaluation not only constitutes a fundamental skill in text summarizing and in related tasks (e.g., underlining, note taking, concept extraction, etc.) but, more generally, it is a prerequisite for any text understanding process.

In recent years we have been working at a new approach to importance evaluation (Fum, Guida, and Tasso, 1985a and 1985b) that is supported by the development of SUSY, an experimental system in the specific domain of descriptive text summarization. Most of the research carried out in this field has been aimed at providing a procedural definition of the concept of importance relying on both structural and semantic knowledge. In this paper we focus on how it is possible to take into account, in evaluating importance, the goals of the reader, in order to investigate how they may influence the evaluation process and its output. In fact, it is expected not only that different representations of the same text will be produced, but also that goals will directly affect the way importance is evaluated.

BASIC ARCHITECTURE OF THE EVALUATOR

The *importance evaluator* is one of the fundamental subsystems of SUSY, and it is specifically devoted to the task of ranking different parts of a text according to their importance (Fum, Guida, and Tasso 1985b).

Several reasons have supported the choice of implementing the importance evaluator by means of a rule-based approach (Waterman and Hayes-Roth, 1978). First of all, the multiplicity and heterogeneity of the knowledge involved in the process of importance evaluation has to be mentioned: linguistic knowledge (both structural and semantic), world knowledge (including both common sense and domain specific knowledge), knowledge about reader's goals, meta-knowledge about how to use linguistic knowledge, world knowledge, and goals in the process of importance evaluation. Second, the concept of importance seems to escape a simple, explicit, algorithmic definition. A conceptual unit of a text can be considered important, for example, because it helps understanding discourse coherence, or because it relates to the discourse topic or topic-focus articulation of discourse, or because it refers to semantically important concepts in the subject domain, or, finally, because it refers to a given reader's goal. A rule-based approach comprising a set of rules that can assign relative importance values to the different conceptual units of a text seems therefore more viable than a traditional deterministic solution, as it can supply all the conceptual and computational tools needed for taking into account in a flexible and natural way the variety of knowledge sources and processing activities that are involved in importance evaluation.
The overall architecture of the evaluator is shown in Figure 1.

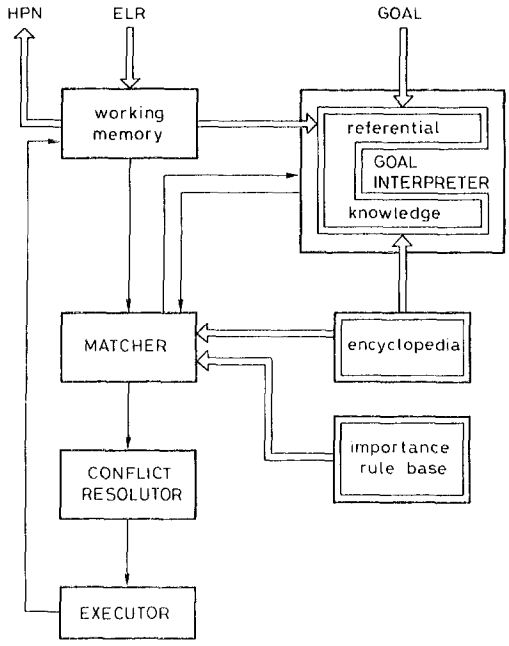

Figure 1. Basic Architecture of the Evaluator.

Basically, it is constituted by the standard modules of a rule-based system (Waterman and Hayes-Roth, 1978) with the addition of a specialized module, namely the *goal interpreter*, devoted to take into consideration the reader's goal.

The evaluator receives in input the internal representation of a natural language text (supplied by another SUSY subsystem, namely the parser) expressed in the ELR *(Extended Linear Representation)* formalism (Fum, Guida, and Tasso, 1984), and an explicit representation of a *goal* to be taken into account for importance evaluation. It produces in output a new

(*) also with: Dipartimento dell'Educazione, Universita' di Trieste, Italy.
(+) also with: Progetto di Intelligenza Artificiale, Dipartimento di Elettronica, Politecnico di Milano, Italy.
(˜) also with: CISM - International Center for Mechanical Sciences, Udine, Italy.

representation called HPN *(Hierarchical Propositional Network),* where integer *importance values* are assigned to the basic conceptual units of the ELR (concepts and propositions), in such a way as to account for the different importance of the constituents of the text.

Two main knowledge bases are available to the evaluator:
- the *importance rule base,* that contains knowledge on the mechanisms that are supposed to be used by human readers in evaluating importance, expressed through IF-THEN rules;
- the *encyclopedia,* that contains specific world knowledge on the subject domain (mostly of structured, taxonomic, descriptive nature), represented through a network of frames.

The IF-part of a rule contains conditions that are evaluated with respect to the current HPN (initially the ELR) contained in the working memory. The THEN-part specifies either an importance evaluation action or an action to be performed to further the analysis (e.g., a strategic choice concerning rule activation, a criterion to solve conflicting evaluations, the activation of a frame of the encyclopedia, etc.). Both parts of a rule may refer to frames of the encyclopedia.
The importance evaluation action contained in the THEN-part of a rule takes usually the form of an assignment of an importance value to a concept or proposition of the ELR. Such an assignment may be absolute (e.g., w(X)=9) or relative (e.g. w(X)=w(X)-3). All successive assignments to a given concept or proposition are not directly executed, but they are stored in a list together with the number of the rules from which they originate, and only at the end of the importance evaluation activity they are globally considered in order to obtain a unique importance value.

The importance rule base includes several classes of rules which account for the different skills used in importance evaluation. Namely (Fum, Guida, and Tasso, 1985b):
- *referential-structural (RS) rules* can derive importance values from the structure of references among conceptual units of the text,
- *rhetoric-structural (TS) rules* derive importance relation from rhetoric predicates of the ELR,
- *structural-semantic (SS) rules* rely on the analysis of specific structural features of the text that have a definite semantic role, such as ISA relations and macro-predicates of the ELR,
- *semantic-encyclopedic (SE) rules* refer to world knowledge contained in the encyclopedia,
- *explicit evaluation (EE) rules* take into account explicit statements concerning importance sometimes purposedly inserted in the text by the author, and, finally,
- *metarules (MT)* embody strategic knowledge that concerns reasoning about importance rules and their use.

The *encyclopedia* is the second knowledge source employed by the evaluator and it contains domain specific knowledge in form of frames. The frames of the encyclopedia embody, in addition to a *header* two kinds of slots:
- *knowledge slots,* that contain domain specific knowledge, represented in a form homogeneous with the ELR language;
- *reference slots,* containing pointers to other frames that deal with related topics in the subject domain.
The operation of the evaluator obeys the basic recognize-act cycle of a rule-based system. More specifically, it is controlled by a forward chaining mechanism which continually updates the working memory, thus transforming it in the final HPN form. The matcher is responsible for recognizing ELR patterns in the working memory which satisfy the IF-part of importance rules. The IF-part of a rule may contain a specific reference to the goal interpreter when it is needed to evaluate the relevance of a given concept or proposition (belonging to the ELR or to the encyclopedia) to the current goal. In such a case the matcher resorts to the goal interpreter, which is able to compute the required *relevance degree,* expressed through a real value in the range (0,1).
When the conflict set has been identified, the conflict resolutor selects the unique rule whose THEN-part will later be executed. System operation ends when the conflict set is empty, i.e. when all available resources for importance evaluation have been used. Several strategies are utilized for performing conflict resolution, and they basically obey two paradigms, namely refraction and ordering (Brownston, Farrel, Kant, and Martin, 1985).

Refraction implies that a rule can not be executed more than once on the same data. On the contrary, it has to be noticed that a single rule can be fired several times on different data during the process of importance evaluation. In fact, the ELR can possibly contain several instances of the patterns conforming to the specific criteria of importance evaluation captured by a rule.
Rule ordering implies that each importance rule is attached a *weight* (an integer value) in such a way as to define an ordering relation among rules. During conflict resolution this ordering is used in two different ways: for selecting the rule with the highest weight, or for discarding rules below a given threshold. Weights are initially assigned statically and are later updated dynamically at run-time. The static ordering is provided when a rule is created, encoding in such a way general selection criteria regarding the priority of using some rules rather than others at the beginning of the importance evaluation process. Dynamic updating of weights allows later on the evaluator to conform to different conflict resolution criteria, adapting its behavior to the actual course of the step-wise transformation of the ELR into the HPN. Null weights are utilized to prevent the possible unwanted execution of a rule. Each time the evaluator is utilized on a new text, the weights are reset to their original static values.

ROLE AND REPRESENTATION OF GOALS

It is apparent that reader's goals have a major role in evaluating importance of a written text. Goals that exist a-priori in the reader's mind, i.e. before reading a text, can affect his judgemental activity in two quite distinct ways. First, the existence of goals can trigger an evaluation mechanism that tends to identify as important those parts of the text which are relevant to the current goal *(goal-directed evaluation).* As goal-directed evaluation strategies coexist with other strategies which are independent of the existence of goals, it is necessary that they would appropriately fit together in such a way as to achieve a correct balance between goal-dependent and goal-independent judgements. Second, goals can have a major role in directing the retrieval and use of encyclopedic knowledge relevant to the current importance evaluation activity *(selective focusing).* In fact, a human reader generally utilizes a lot of specific world knowledge when evaluating the importance of a text, and the reminding from long term memory of the pieces of knowledge to be used in a given context is often triggered by his a-priori goals. In this case, goals do not directly contribute to the importance evaluation process, but can affect it in an indirect way through identification of pertinent world knowledge to be used by other goal-independent importance evaluation strategies.

Using goals in importance evaluation poses two classes of problems to the system designer:
- How to represent goals?
- How to match goals with pieces of the ELR or encyclopedia for implementing the mechanisms of goal-directed evaluation and selective focusing?
The former of the above points will be dealt with in the sequel of this section, while the latter will be the subject of the next section.
Several kinds of reader's goals are possible with respect to their generality, level of abstraction, articulation of content, richness of details, etc. It is apparent that goals, according to their different nature, may range from a light emphasis of the reader's intentions to a quite specific query. More precise and articulated the goals are, more focused is the attention on well defined and specific objects, and, accordingly, goal-dependent importance evaluation strategies become more appropriate and useful. Moreover, as goals become more and more specific and rich, importance evaluation tends to mingle with information retrieval and question-answering.

As a basic design choice, we restrict our attention (at least for the moment) to classes of goals which are reasonably general and simple (but not necessarily explicit, univocal, clear, or easy to interpret!), in such a way as to keep focusing on importance evaluation without intermixing too much our model with different issues, such as information retrieval and question-answering.
This decision has heavy implications on the design of the *goal representation language.* Although in principle nothing less than the full ELR formalism should be used for expressing goals, we restrict our attention to a largely simplified subset of it.
Let us consider a *goal vocabulary* (GV) containing a collection of keywords relevant to a given subject domain, and assume as an adequate

representation for a goal a propositional expression over GV made up using *and, or,* and *not* connectives. Note that the goal vocabulary GV may be redundant, i.e. it may contain several words that refer to partially overlapping concepts. Also it is implicitly assumed that the general topic of discourse is fixed and always tacitly understood: words of GV only specify a facet, a viewpoint or a detail of interest, but they can not change or modify (e.g., through limitations or specifications) the topic of discourse. Moreover, it is assumed that the size of GV can be kept reasonably small, although each time a new interesting concept has to be included in the coverage of the goal representation language, GV must be enlarged accordingly.

INTERPRETING GOALS

After having introduced a representation language to be used for specifying goals, we tackle in this section the problem of how it is possible to obtain goal dependent importance evaluation.

The first possibility that comes to mind consists in labeling a-priori each frame of the encyclopedia with words of the goal vocabulary GV and, then, to match words used for expressing goals with such labels, taking appropriately into account the logical connectives *and, or,* and *not.* This solution is only possible for the encyclopedia, as it is quite impossible to label a-priori unknown pieces of ELR. It shares some basic features with the approach proposed by DeJong (1979) that assumes an a-priori definition of the concept of importance coded into an appropriate set of scripts. This solution, however, has several shortcomings:
(a) it is rigid, as any change in GV necessitates that the labeling of the encyclopedia is changed accordingly;
(b) it hiddens the reasons and criteria adopted for labeling the frames of the encyclopedia, thus preventing any further use of this information (for example, in generating justifications of the evaluation produced);
(c) it makes the encyclopedia heavily dependent on the specific use of importance evaluation and on the particular goal vocabulary currently considered.

A better solution, that can cover both cases of comparison with the encyclopedia and ELR and does not require any preliminary labeling procedure, is direct matching of words of GV appearing in the goal specification with frames of the encyclopedia or pieces of the ELR, taking appropriately into account the meaning of logical connectives. This possibility, however, would also be largely unsatisfactory. It does not allow taking into account, for example, the diversity of terminology, i.e. the fact that the same concept may be referred to by means of different words in the goal specification and in the piece of knowledge to be matched; it does not allow dealing with concepts at different levels of abstraction, and, more importantly, it does not allow expressing different degrees of relevance.

The above analysis of inadequacies of some preliminary design proposals allows stating the following requirements for the goal interpreter:
- it should allow to keep the goal vocabulary GV and the encyclopedia independent from each other, in such a way as changes in the former do not affect the latter, and they can be designed and updated separately;
- it should support an explicit representation of the conceptual connection between the goal vocabulary GV and the encyclopedia and ELR, in such a way as the role of goals in importance evaluation can be easily controlled by the system designer and, if necessary, explained and justified to the user;
- it should allow dealing with diversity of terminology, expression, context, and level of abstraction;
- it should allow dealing with a full range of relevance degrees.

We propose here a first step towards the design of a goal interpreter satisfying the above mentioned requirements. Such an interpreter takes in input a goal specification, expressed in the goal representation language, and a fragment of ELR taken from the internal representation of the text or from the frames of the encyclopedia. Its task is to compute the *relevance degree* of the ELR fragment according to the given goal. To this purpose, the goal interpreter utilizes a *referential knowledge* base, i.e., a semantic network whose nodes are either *atomic concepts* that represent basic items in the subject domain or *definitional concepts,* i.e., structures that are used in order to define the meaning of atomic concepts. The arcs of the network connect pairs of nodes linked together by some conceptual relation such as synonymy, antonymy, generalization, specification, definition, attribute, etc. Each arc is tagged by a label, indicating the conceptual relation linking the two concepts, and by a real number in the range (0 - 1) which represents the *relation degree* that characterizes the link between the two concepts. The referential knowledge base represents the main knowledge source utilized by the goal interpreter in evaluating the relevance degree of an ELR fragment to a given goal. General knowledge regarding the subject domain and, more specifically, concerning the discourse topic is thus wired in the referential knowledge base.

The semantic network which constitutes the referential knowledge base is accessed starting from the ELR fragment in parallel. At this point a bidirectional search process, aimed at finding a path connecting the two entry points, begins. The search is complicated by the fact that some nodes of the network are constituted by atomic concepts whereas other nodes are definitional. It is possible to proceed from a definitional node onwards if and only if all the concepts constituting the definition can be matched, directly or through intermediate nodes, with ELR expressions. The search process terminates when a path connecting the goal and the ELR fragment is found. An appropriate function taking into account the relation degrees of the arcs in the path is computed and the result represents the relevance degree of the ELR fragment to the given goal. Whenever possible an optimum path (i.e. a path with the highest relevance degree) should be looked for, but such a search generally poses hard problems from a computational point of view.

AN EXAMPLE

This section is devoted to present an example of operation of the goal interpreter. Let us consider the following fragment of text, taken from Christian (1983: 11):

> *"... The UNIX system is a moderately complex operating system. It is far simpler than the operating systems that run on maxicomputers, but it has much more capabilities than most operating systems that run on microcomputers. For example, the UNIX system allows several programs to run simultaneously. ..."*

The purpose of this example is to show how the goal interpreter is able to identify that the last sentence of the text is to be considered important if evaluated with reference to the goal USE. By applying usual referential-structural rules, the concept UNIX is stressed as important since it is highly referenced in the text. The ELR representation of the last sentence is:

180 ALLOW (UNIX, 190, P)
190 RUN (VV3)
200 *PROGRAM (VV3)
210 SEVERAL (VV3)
220 SIMULTANEOUSLY (190, P)

Consider now the following semantic-encyclopedic rule:

Rule SE26

IF there is a proposition P A(X,Y) such that:
- A ISA PERFORM
- w(X) >= high
- the relevance degree of Y to the current goal is >= 0.5

THEN set w(P) = w(X).

The rationale behind rule SE26 is that a sentence concerning an important concept is also considered important when its predicate is of kind PERFORM and its second argument (i.e., what is predicated about the important concept) is relevant to the current goal.
The first two clauses of the IF-part of the rule match proposition 180, since ALLOW ISA PERFORM and w(UNIX)=high. For what concerns the third clause, a deeper analysis involving the goal interpreter is needed.

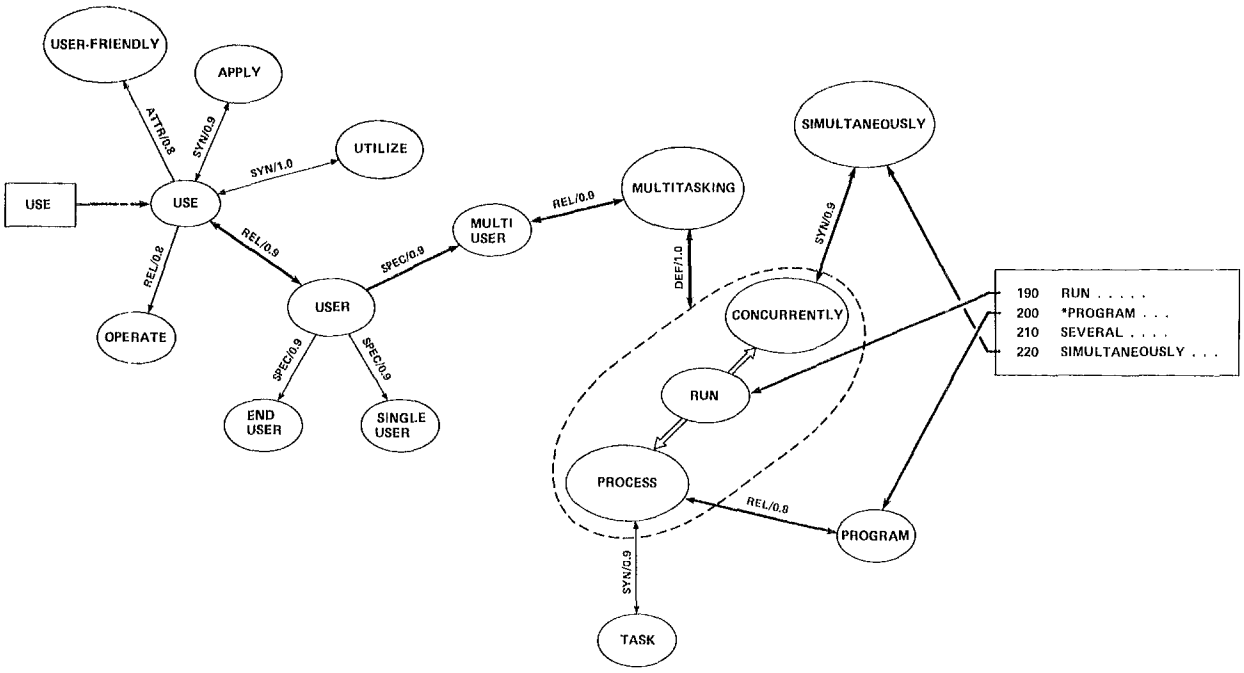

Figure 2. Referential Knowledge Utilized by the Interpreter.

More specifically, the relevance degree of proposition 190, which in turn involves also propositions 200, 210, and 220, has to be evaluated with reference to the goal USE. The portion of referential knowledge utilized by the goal interpreter in this specific case is shown in Figure 2.
The network is entered through the word USE, corresponding to the goal, and the nodes RUN, PROGRAM, and SIMULTANEOUSLY. By moving through the network from both entries, the path drawn in bold lines in Figure 2 is identified. The definitional node corresponding to the MULTITASKING concept is entered from the ELR through multiple (namely, three) arcs. The overall relevance degree of the path is computed by multiplying the relation degrees of its arcs and the result 0.58 is obtained. It should be noted that, among the several arcs entering a definitional node, only that with the lowest relation degree is considered for the computation. In this way, rule SE26 can be applied and, consequently, the importance value of proposition 180 is set to high.

CONCLUSIONS

The evaluator described in the paper is presently running in a prototype version (with about 50 rules and a small encyclopedia of about 40 frames) written in LISP on a SUN-2 workstation. It has been extensively tested on selected cases (extracted from textbooks on operating systems), and it includes a preliminary implementation of the goal interpreter based on a simplified version of the referential knowledge base. Research work devoted to extend and refine this first version of the goal interpreter and to test its performance is now ongoing.

Several open problems and challenging topics will be the subject of future research. Among these we mention:
- extending the goal representation language;
- merging the encyclopedia and the referential knowledge base into a unified structure that encompasses all knowledge available on the subject domain and can make it available in an effective way to the relevant modules of the evaluator (Sowa, 1984);
- considering goals that can change during text understanding, taking into account the topic-focus articulation of discourse - (Hajicova' and Sgall, 1984);
- developing an explanation module that can justify the reasons behind the evaluations produced by the system.

REFERENCES

Brownston L., Farrel R., Kant E., and Martin N. (1985). *Programming Expert Systems in OPS5.* Reading, MA: Addison-Wesley.

Christian K. (1983). *The Unix Operating System.* New York, NY: Wiley.

DeJong G.F. (1979). Skimming Stories in Real Time: An experiment in integrated understanding. *Research Report #158,* Yale University Department of Computer Science, New Haven, CT.

Fum D., Guida G., and Tasso C. (1984). A Propositional Language for Text Representation. In B.G. Bara and G. Guida (Eds.), *Computational Models of Natural Language Processing,* Amsterdam, NL: North-Holland, 121-163.

Fum D., Guida G., and Tasso C. (1985a). A Rule-Based Approach to Evaluating Importance in Descriptive Texts. *Proc. 2nd Conf. of the European Chapter of the Association for Computational Linguistics,* Geneva, Switzerland, 244-250.

Fum D., Guida G., and Tasso C. (1985b). Evaluating Importance: A step towards text summarization. *Proc. 9th Int. Joint Conf. on Artificial Intelligence,* Los Angeles, CA, 840-844.

Hajicova' E. and Sgall P. (1984). From Topic and Focus of a Sentence to Linking in a Text. In B.G. Bara and G. Guida (Eds.), *Computational Models of Natural Language Processing,* Amsterdam, NL: North-Holland, 151-163.

Sowa, J.F. (1984). *Conceptual Structures* Reading, MA: Addison-Wesley.

Waterman D.A. and Hayes-Roth F. (Eds.) (1978). *Pattern-Directed Inference Systems.* New York, NY: Academic Press.