Towards Interactive Text Understanding

Marc Dymetman* **Aurélien Max*+** **Kenji Yamada***
(*) Xerox Research Centre Europe, Grenoble
(+) CLIPS-GETA, Université Joseph Fourier, Grenoble
{marc.dymetman,aurelien.max,kenji.yamada@xrce.xerox.com}

Abstract

This position paper argues for an *interactive* approach to text understanding. The proposed model extends an existing semantics-based text authoring system by using the input text as a source of information to assist the user in *re-authoring* its content. The approach permits a reliable deep semantic analysis by combining automatic information extraction with a minimal amount of human intervention.

1 Introduction

Answering emails sent to a company by its customers — to take just one example among many similar text-processing tasks — requires a reliable understanding of the content of incoming messages. This understanding can currently only be done by humans, and represents the main bottleneck to a complete automation of the processing chain: other aspects could be delegated to such procedures as database requests and text generation. Current technology in natural language understanding or in information extraction is not at a stage where the understanding task can be accomplished reliably without human intervention.

In this paper, which aims at proposing a fresh outlook on the problem of text understanding rather than at describing a completed implementation, we advocate an interactive approach where:

1. The building of the semantic representation is under the control of a human author;

2. In order to build the semantic representation, the author interacts with an intuitive textual interface to that representation (obtained from it through an NLG process), where some "active" regions of the text are associated with menus that display a number of semantic choices for incrementing the representation;

3. The raw input text to be analyzed serves as a source of information to the authoring system and permits to associate likelihood levels with the various authoring choices; in each menu the choices are then ranked according to their likelihood, allowing a speedier selection by the author; when the likelihood of a choice exceeds a certain threshold, this choice is performed automatically by the system (but in a way that remains revisable by the author).

4. The system acts as a flexible understanding aid to the human operator: by tuning the threshold at a low level, it can be used as a purely automatic, but somewhat unreliable, information extraction or understanding system; by tuning the threshold higher, it can be used as a powerful interactive guide to building a semantic interpretation, with the advantage of a plain textual interface to that representation that is easily accessible to general users.

The paper is organized as follows. In section 2, we present a document authoring system, MDA, where the author constructs an internal semantic representation, but interacts with a textual realization of that representation. In section 3, we explain how such a system may be extended into an Interactive Text Understanding (ITU) aid. A raw input document acts as an information source that serves to rank the choices proposed to the author according to their likelihood of "accounting" for information present in the input document. In section 4, we present current work on using MDA for legacy-document normalization and show that this work can provide a first approach to an ITU implementation. In section 5, we indicate some links between these ideas and current work on interactive statistical MT (TransType), showing directions towards more efficient implementations of ITU.

2 MDA: A semantics-based document authoring system

The MDA (Multilingual Document Authoring) system [Brun et al 2000] is an instance (descended from Ranta's Grammatical Framework [Ranta 2002]) of a text-mediated interactive natural language generation system, a notion introduced by [Power and Scott 1998] under the name of WYSIWYM. In such systems, an author

gradually constructs a semantic representation, but rather than accessing the evolving representation directly, she actually interacts with a natural language text generated from the representation; some regions of the text are active, and correspond to still unspecified parts of the representation; they are associated with menus presenting collections of choices for extending the semantic representation; the choices are semantically explicit and the resulting representation contains no ambiguities. The author thus has the feeling of only interacting with text, while in fact she is building a formal semantic object. One application of this approach is in multilingual authoring: the author interacts with a text in her own language, but the internal representation can be used to generate reliable translations in other languages. Fig. 1 gives an overview of the MDA architecture and Fig. 2 is a screenshot of the MDA interface.

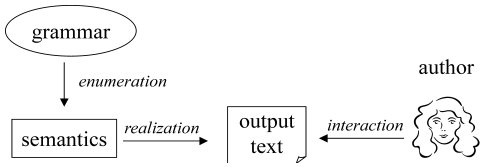

Fig. 1: Authoring in MDA. A "semantic grammar" defines an enumerable collection of well-formed partial semantic structures, from which an output text containing active regions is generated, with which the author interacts.

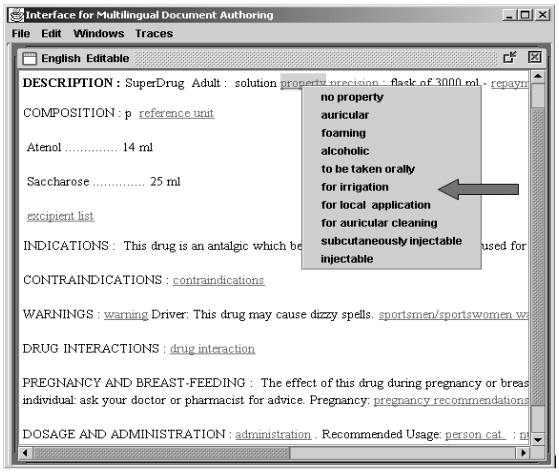

Fig. 2: Snapshot of the MDA system applied to the authoring of drug leaflets.

3 Interactive Text Understanding

In the current MDA system, menu choices are ordered statically once and for all in the semantic grammar[1]. However, consider the situation of an author producing a certain text while using some input document as an informal reference source. It would be quite natural to assume that the authoring system could use this document as a source of information in order to prime some of the menu choices.

Thus, when authoring the description of a pharmaceutical drug, the presence in the input document of the words *tablet* and *solution* could serve to highlight corresponding choices in the menu corresponding to the pharmaceutical form of the drug. This would be relatively simple to do, but one could go further: rank menu choices and assign them confidence weights according to textual and contextual hints found in the input document. When the confidence is sufficiently high, the choice could then be performed automatically by the authoring system, which would produce a new portion of the output text, with the author retaining the ability of accepting or rejecting the system's suggestion. In case the confidence is not high enough, the author's choice would still be sped up through displaying the most likely choices on top of the menu list.

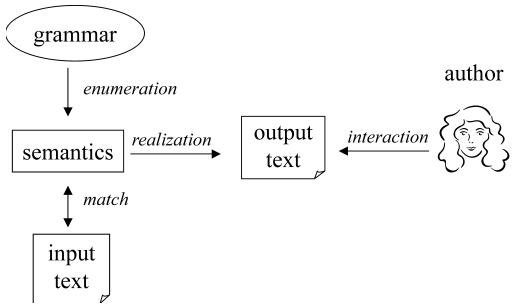

Fig. 3: Interactive Text Understanding.

This kind of functionality is what we call a *text-mediated interactive text understanding system*, or for short, an *ITU* system (see Fig. 3).[2]

[1] While the order between choices listed in a menu does not vary, certain choices may be filtered out depending on the current authoring context; this mechanism relies on unification constraints in the semantic grammar.

[2] Note that we do not demand that the semantic representation built with an ITU system be a complete representation of the input document, rather it can be a structured description of some thematic aspects of that document. Similarly, it is OK for the input document not to contain enough information permitting the system or even the author to "answer" certain menus: then some active regions of the output text remain unspecified.

We will now consider some directions to implement an ITU system.

4 From document normalization to ITU

A first route towards achieving an ITU system is through an extension of ongoing work on document normalization [Max and Dymetman 2002, Max 2003]. The departure point is the following. Assume an MDA system is available for authoring a certain type of documents (for instance a certain class of drug leaflets), and suppose one is presented a "legacy" document of the same type, that is, a document containing the same type of information, but produced independently of the MDA system; using the system, a human could attempt to "re-author" the content of the input legacy document, thus obtaining a normalized version of it, as well as an associated semantic representation.

An attempt to automate the re-authoring process works as follows. Consider the virtual space of semantic representations enumerated by the MDA grammar. For each such representation, produce, through the standard MDA realization process[3] a certain more or less rough "descriptor" of what the input text should contain if its content should correspond to that semantic representation; then define a similarity measure between this descriptor and the input text; finally perform an admissible heuristic search [Nilsson 1998] of the virtual space to find the semantics whose descriptor has the best similarity with the input text. This architecture can accomodate more or less sophisticated descriptors: from bags of content-words to be intersected with the input text, up to predicted "top-down" predicate-argument tuples to be matched with "bottom-up" tuples extracted from the input text through a rough information-extraction process.

Up to now the emphasis of this work has been more on automatic reconstruction of a legacy document than on interaction, but we have recently started to think about adapting the approach to ITU. The heuristic search that we mentioned above associates with a menu choice an estimate of the best similarity score that could be obtained by some complete semantic structure extending that choice. It is then possible to rank choices according to that heuristic estimate (or some refinement of it obtained by deepening the search a few steps down the line), and then to propose to the author a re-ranked menu.

While we are currently pursuing this promising line of research because of its conceptual and algorithmic simplicity, it has some weaknesses. It relies on similarity scores between an input text and a descriptor that are defined in a somewhat *ad hoc* manner, it depends on parameters that are fixed *a priori* rather than by training, and it is difficult to associate with confidence levels having a clear interpretation.

A way of solving these problems is to move towards a more probabilistic approach that combines advantages of being built on accepted principles and of having a well-developed learning theory. We finally turn our attention to existing work in this area that holds promise for improving ITU.

5 Towards statistical ITU

Recent research on the interactive statistical machine translation system TransType [Foster et al, 1997; Foster et al, 2002] holds special interest in relation to ITU. This system, outlined in Fig. 4, aims at helping a translator type her (unconstrained) translation of a source text by predicting sequences of characters that are likely to follow already typed characters in the target text; this prediction is done on the basis of information present in the source text. The approach is similar to standard statistical MT[4], but instead of producing one single best translation, the system ranks several completion proposals according to a probabilistic confidence measure and uses this measure to optimize the length of completions proposed to the translator for validation. Evaluations of the first version of TransType have already shown significant gains in terms of the number of keystrokes needed for producing a translation, and work is continuing for making the approach effective in real translation environments.

If we now compare Fig. 3 and Fig. 4, we see strong parallels between TransType and ITU: language model enumerating word sequences *vs*

[3] Which was initially designed to produce parallel texts in several languages, but can be easily adapted to the production of non-textual "renderings" of the semantic representations.

[4] Initially statistical MT used a noisy-channel approach [Brown et al. 1993]; but recently [Och and Ney 2002] have introduced a more general framework based on the maximum-entropy principle, which shows nice prospects in terms of flexibility and learnability. An interesting research thread is to use more linguistic structure in a statistical translation model [Yamada and Knight 2001], which has some relevance to ITU since we need to handle structured semantic data.

grammar enumerating semantic structures, source text *vs* input text as information sources, match between source text and target text *vs* match between input text and semantic structure. In TransType the interaction is directly with the target text, while in ITU the interaction with the semantic structure is mediated through an output text realization of that structure. We can thus hope to bring some of the techniques developed for TransType to ITU, but let us note that some of the challenges are different: for instance training the semantic grammars in ITU cannot be done on a directly observable corpus of texts.[5]

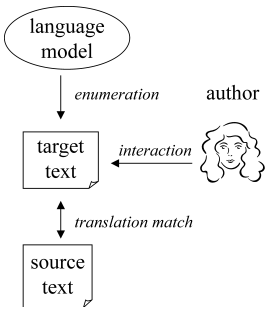

Fig. 4: TransType.

6 Conclusion

We have introduced an interactive approach to text understanding, based on an extension to the MDA document authoring system. ITU at this point is more a research program than a completed realization. However we think it represents an exciting direction towards permitting a reliable deep semantic analysis of input documents by complementing automatic information extraction with a minimal amount of human intervention for those aspects of understanding that presently resist automation.

[5] Let us briefly mention that we are not the first to note formal connections between natural language understanding and statistical MT. Thus, [Epstein 1996], working in a non-interactive framework, draws the following parallel between the two tasks: while in MT, the aim is to produce a target text from a source text, in NLU, the aim is to produce a semantic representation from an input text. He then goes on to adapt the conventional noisy channel MT model of [Brown et al 1993] to NLU, where extracting a semantic representation from an input text corresponds to finding: argmax(Sem) {p(Input|Sem) p(Sem)}, where p(Sem) is a model for generating semantic representations, and p(Input|Sem) is a model for the relation between semantic representations and corresponding texts. See also [Berger and Lafferty 1999] and [Knight and Marcu 2002] for parallels between statistical MT and Information Retrieval and Summarization respectively. On a different plane, in the context of interactive NLG, [Nickerson 2003] has recently proposed to rank semantic choices according to probabilities estimated from a corpus; but here the purpose is not text understanding, but improving the speed of authoring a new document from scratch.

Acknowledgements

Thanks for discussions and advice to C. Boitet, C. Brun, E. Fanchon, E. Gaussier, P. Isabelle, G. Lapalme, V. Lux and S. Pogodalla.

References

[Berger and Lafferty 1999] Information Retrieval as Statistical Translation, SIGIR-99

[Brown, Della Pietra, Della Pietra and Mercer 1993] The Mathematics of Statistical Machine Translation: Parameter Estimation. Computational Linguistics 19(2), 1993

[Brun, Dymetman and Lux 2000]. Document Structure and Multilingual Text Authoring, INLG-2000

[Epstein 1996] Statistical Source Channel Models for Natural Language Understanding, PhD Thesis, New York University, 1996.

[Foster, Isabelle and Plamondon, 1997] Target-Text Mediated Interactive Machine Translation, Machine Translation, 12:1-2, 175-194, Dordrecht, Kluwer, 1997.

[Foster, Langlais and Lapalme, 2002] User-Friendly Text Prediction for Translators, EMNLP-02

[Knight and Marcu 2002] Summarization beyond sentence extraction: A Probabilistic Approach to Sentence Compression, Artificial Intelligence, 139(1), 2002.

[Max and Dymetman 2002] Document Content Analysis through Fuzzy Inverted Generation, in AAAI 2002 Spring Symposium on Using (and Acquiring) Linguistic (and World) Knowledge for Information Access, 2002

[Max 2003]. Reversing Controlled Document Authoring to Normalize Documents. In the proceedings of the EACL-03 Student Research Workshop, 2003

[Nickerson 2003]. Statistical Models for Organizing Semantic Options in Knowledge Editing Interfaces. In AAAI Spring Symposium workshop on natural language generation in spoken and written dialogue, 2003.

[Nilsson 1998] Artificial Intelligence: a New Synthesis. Morgan Kaufmann, 1998.

[Och and Ney 2002] Discriminative Training and Maximum Entropy Models for Statistical Machine Translation, ACL02

[Power and Scott 1998] Multilingual Authoring using Feedback Texts. COLING/ACL-98.

[Ranta 2002] Grammatical Framework: A Type-Theoretical Grammar Formalism, Journal of Functional Programming, September 2002.

[Yamada and Knight 2001] A Syntax-based Translation Model, ACL-01.