

Intentional Structures of Documents

Tazi Saïd

Laboratoire d'Interaction Homme Systèmes
(LIHS)
Université de Toulouse1,
Place A. France
F-31042 Toulouse Cedex
tazi@cict.fr

Fabrice Evrard

Laboratoire Informatique et Mathématiques
Appliquées
(GRAAL - LIMA - IRIT - UMR 5505)
ENSEEIH 2, Rue Camichel
F-31071 Toulouse Cedex
Fabrice.Evrard@enseeiht.fr

ABSTRACT

Document structures constitute a means to organize document parts in terms of logical elements (e.g. headings, paragraphs, sections, etc.). Document structures can be looked at in terms of layout and formatting features such as pages, columns, and so on. Despite the wide diversity of document structures, authors intentions are not taken into account. The concept of intentions is considered here as the effects that authors intend to have on their readers. It includes the reasons that led the authors to write the document and to select certain features instead of others. Intentional structures are a set of relations between document fragments that make explicit authors' intentions. Intentional structures are generally implicit. We want to help authors to make some of their intentions more explicit. This paper presents a new model based on speech act theory dealing with intentional structures and arguments of how it can be used to enhance written communication.

KEYWORDS: Document structures, Intentional structures, Writing Acts

INTRODUCTION

Authoring systems do not offer the possibility to explicit the intentional structures of documents. Despite the wide diversity in the concept of document structure, the authors' intentions are not taken into account. By intention, we mean what effects the writers intend to have on their readers, why they use one means instead of another to carry out the writing task, the rationale of tasks performed on the document, and so on. Intentional structures can be expressed as a set of relations between document fragments. Each fragment is related to what we call intentions. Intentional structures are generally implicit, and sometimes the writers are unaware of them. We hypothesize that if authoring systems could allow authors to take this notion into account, they could contribute to rendering a text more explicit to readers.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
HT'01 8/01 Aarhus, Denmark
© 2001 ACM ISBN 1-59113-420-7/01/0008...\$5.00

The model we propose here to deal with this notion is based on speech act theory. This theoretical aspect will be presented in the next section. A discussion of how this model can be used will be presented in the last section.

THEORETICAL BACKGROUND

According to speech acts theory, the concept of act can be applied not only to spoken interaction but to written communication as well, see [1] and [4]. An author like a speaker, will have a set of effects that he or she wishes to produce in the intended recipient.

Hypertext links are made deliberately by authors to add extra-information to texts or to structure parts of the document. Identifying and representing the illocutionary acts of the process of writing and linking passages will help authors to record reasons to create links in hypertexts, and will likewise help readers to find out the intentions of writing and linking. This help will be made explicit in an authoring system as a set of commands that the authors want to perform. For example, instead of simply linking two passages that are an explanation and the text being explained, the author could perform an "Explain command" that links the two passages.

Domain Acts and Meta Acts

The Speech act theory distinguishes between acts and meta acts [4], [2]. Acts represent the intention of the utterance. Meta act models use the notion of act to account for communication about communication. The Intentional Structure (I.S.) model we propose is based on this theory. Each I.S. is considered as a set of predicates that represent both meta acts and domain acts that lead to perform the authors' intentions. Meta acts are verbs that express the conventions used to carry out the writing. The meta acts that we have considered are : formatting meta acts, document-organizing meta acts, and linking meta acts. [5]

Intentional Structures

The intentions we are interested in are those which the author has in mind when he is planning, drafting, writing, and revising a document. We suppose that all authors ask themselves at least the following questions which we will

call writing questions : “Who is the addressee in this part of the document?”, “What does he/she (the author) want to say by writing it?”, “What does he/she want to achieve by writing that?”, “Why does he/she want to do (or say) it?”, “How does he formulate the statement?”, “Why does he/she do it in this way rather than in another one?”. We consider that the efficiency of written communication depends on the answers that authors give to their writing questions. The model we developed is a set of predicates that represent the answers to the writing questions.

When an author performs a written communication he/she achieves a set of acts that express simultaneously the writing action itself, the goal of this action, an act which relates to the writing, that we call “Means” and an act that can express the reason for this action. The intention also depends on the recipient of the communication. We have adopted the following generic form to represent an intention:

Intention(Agent1, Agent2, Action, Goal, Means, Reason)

Where :

- Agent1* is the author of the action. It might be the writer, or a reader in the case of revision, or annotation;
- Agent2* is the agent for whom the action is intended, generally it is the reader;
- Action* is an act that expresses what the author of the intention wants to carry out;
- Goal* is an act that expresses what the author wants to do by performing the action;
- Means* is an act that expresses the type of action achieved on the text; it is a meta act ;
- Reason* is an act that expresses why the author carries out the action.

Sample

Let us suppose that a writer intends to carry out a quotation (action) with the aim of documenting his own text (goal). He/she uses a text between inverted commas (Means) to convince the reader (Reason) of the position that he supports. We express this intention by :

Intention(Writer, Reader, To_Cite (Writer, Author, Citation-text), To_Document (Writer,Writer-text, Citation-Text), To_Quote (Writer, Citation-text), Convince (Writer, Reader, Writer-text))	Agent1, Agent2 Action Goal Means Reason
--	---

One recognizes that the Writer and Reader are respectively Agent1 and Agent2. The action is To_Cite: the writer quotes an author by using a Citation-Text, the Goal is To_Document what is being written (Writer-text) by the Citation-text. The means is a quotation (To_Quote), and the reason for it is that the Writer wants to convince the reader about his text.

Several alternatives are possible: the writer could quote, not to convince, but only to inform; he could also document, not by quoting, but by adopting the quotation text, or by using a different feature as typing characters in bold or underlining.

The aim of this model is to take into account all these possibilities, and to increase the writers’ awareness of what he/she is doing, and the effect that the writing might have on his/her readers.

DISCUSSION

The model of I.S. is under development as a conceptual model of @ware, an authoring system under implementation [3]. One of first conclusions of this research is that intentions depend on the domain the author is writing about. The main idea of implementation issues is to consider new commands that let the user to make his/her intentions explicit while writing; the system adds xml tags that allow the storage and the processing of intentional structures. Each fragment of the document must be tagged as an element that expresses the intentional structure it belongs to. This will allow to access the document in queries terms, as for : “What does the author mean here?”, or “Why does he say that?”.

The consequences of making the author’s intentions explicit are numerous and varied. Nevertheless, this requires a great effort from the authors. This problem is due to the fact that habits are hard to change. Over the centuries, mankind has inherited a vast amount of writing traditions. The resolution of such a problem depends on the wealth of accumulated riches of the system and on the users’ versatility. The authoring system must provide a rich set of commands that let the authors make their intentions explicit easily. The system must also let the users change their set of intention commands. Revisions of documents should be based on I.S. For example, instead of the authors typing a string in bold, they may express their intention by clicking on an “*Insist*” command. The system then puts the string in bold if it is the formatting feature for formatting intentions.

ACKNOWLEDGMENTS

Anne Pechou provided astute advice towards improving the quality of this paper.

REFERENCES

1. Austin, J. *How to do things with words*. Cambridge, MA: Harvard University Press, 1962.
2. Carbonell, J. G. (1982). Meta-language utterances in purposive discourse, Technical report CMU-CS-82-185, Department of Computer Science, Carnegie-Mellon University.
3. Evrard F. and Tazi S., *Structures intentionnelles de la communication écrite pour la création interactive de documents* , MFI'01, Modèles Formels pour l'Interaction, Toulouse, 21-23 Mai 2001 , can be accessed at : <http://lihs.univ-tlse1.fr/tazi/Publications/Tazi-Evrard-MFI01.pdf>
4. Searle, J. *Speech acts*. Cambridge: Cambridge University Press, 1969.
5. Tazi, S. and Novick, D., *Actes de la communication écrite*, Proceedings of Ergonomie et Informatique Avancée (Ergo-IA 98), Biarritz, FR, November, 1998., can be accessed at : http://lihs.univ-tlse1.fr/tazi/Publications/ace_ergoia98.html