Generating User-Adapted Hypermedia from Discourse Plans

Berardina De Carolis, Fiorella de Rosis, Sebastiano Pizzutilo

Dipartimento di Informatica, Università di Bari
Via Orabona, 4, 70126, Bari - Italy
tel. +39-80-5443284, fax. +39-80-5443196
{nadja, derosis, pizzutil}@aos2.uniba.it
http://aos2.uniba.it:8080/IntInt.html

Abstract
This paper describes how hypermedia instruction manuals can be generated from a discourse plan: intentional and rhetorical knowledge in the plan tree are employed to construct the hypermedia elements, to assemble them into pages and to introduce links among them. Plan tree transformation functions are defined so as to ensure that generated hypermedia respect interface usability and Web design criteria. User adaptation is performed by parametrizing these functions: this method enables building the hypermedia according to a dynamic assessment of the user needs.

1. Introduction

The generation of complex documents is usually performed in two main steps. In the first of them, discourse planning establishes information items to be introduced in the document and their presentation order. In the second one, surface generation translates the plan into a natural language or multimedia document. Rhetorical Structure Theory (RST: Mann and Thompson,1988) plays a crucial role in the whole generation process (Hovy,1993; Moore and Paris, 1993; De Carolis et al, 1996) by ensuring coherence of the text. In user-adapted presentations, the generation strategy represented in plan operators formalizes what the Speaker believes to be important to mention in the document, as a function of hypotheses about the Addressee's knowledge, preferences and goals. When these hypotheses are incorrect, the generated text reveals to be inadequate: some systems then give the Addressee the opportunity of interacting with the Generator to require more details or to refine the text (see, e.g., Mittal and Moore,1995).

After about ten year experiences of one-shot document generation, research is aimed, more recently, at studying whether methods developed in that context can be extended to the generation of user-adapted hypermedia. In the distribution of tasks between user and system that is proposed in this case, selection of the document subset to view is left to the user, whereas the system is responsible for arranging relevant data in a hypermedia page by organizing them in the appropriate form and language (see, for instance, Milosaljievic and Dale,1996). Methods are derived directly from those employed for one shot texts: the part of the document that the user wishes to explore is generated dynamically in one page, with links to pages with potentially interesting details.

This method is efficient when documents about concepts or objects descriptions have to be generated: most applications developed so far are in this area. In this case, it is reasonable to presume that users know which information they want to look at, and that the system is only left the task of providing this information in an adequate form. However, this idea cannot be generalized to all types of documents: in particular, to instruction manuals about how to perform some action or to use some device. Behind these documents there is, in general, a strong Instructor's intention to orient the Addressees, which is a function of their presumed level of experience. This intention is made explicit, in the document, by means of some appropriate presentation techniques and corresponds to the main communicative goal and to the way that this goal is decomposed into sub-goals.

For instance: in a manual about how to use a domestic blood pressure measurer (which is the case study that we will consider in this paper), the main purpose is to teach how to measure the pressure. To achieve this purpose, one has to teach: (i) how to prepare the device, (ii) how to connect the device to the patient and (iii) how to make the measurement. In order to teach how to prepare the device, one has to teach: (i1) how to switch it on, (i2) how to check that the device is ready for use ...and so on. This description of the goal/subgoals hierarchy is usually mirrored in a hierarchy of document Sections, Subsections and Paragraphs. According to their degree of experience, users can (i) glance non systematically at the manual to look for the Section on which to concentrate their attention, (ii) read it systematically or (iii) read the Index to find out the part of interest to them. In an adaptive hypermedia manual, one should reproduce these document-reading attitudes into different structures, and should activate the most appropriate navigation style according to the user characteristics and to the context of use. Cognitive psychology research is providing evidence for adaptation criteria. For example: 'Index navigation' is preferable for book-like hypertexts whereas 'Page navigation' gives better performance in texts whose structure is a symmetrical hierarchy (Wright and Lickorish, 1989); directed access and guidance mechanisms should be provided in 'exploratory learning' (Hammond and Allison, 1989); browsing and searching information gathering tasks influence preference for linear vs hierarchical or network structures (De Vries and De Jong, 1997), and so on.

In the research described in this paper, we investigated how to generate a user-customized hypermedia instruction manual by studying, in particular, the following issues: which structure should such a manual have? How should this structure be adapted to the user characteristics? Is the discourse plan a suited knowledge source for this generation process? Can RST be employed successfully to this purpose, as it was for one-shot document generation? The paper answers to these questions by describing the methods developed in GeNet, an object-oriented prototype for generating hypermedia manuals on the WWW. In Section 2, we outline the main components of instruction manuals. In Section 3, we describe knowledge embodied in a discourse plan and how it is represented (Section 3.1), the hypermedia structure and generation criteria (Section 3.2) and the algorithm (Section 3.3). User adaptation methods are illustrated in Section 4. Some evaluation about results obtained so far is finally given in Section 5.

2. The case study

An instruction manual usually includes two main sections:

a. a ***descriptive*** section illustrates *device elements*, with the function performed by each of them, and *technical terms* employed.

b. an ***instructional*** section describes how to perform the action or to use the device.

In the second section, reference is frequently made to items described in the first one: separation between the two components increases the difficulty with which instructions can be followed, especially when the device structure and the role of each item are not clear. A deep integration between the two components therefore favours understanding the instruction process. This integration can be achieved by implementing the manual as a user adapted hypermedia. In this hypothesis, each component is described in a separate section, that the user can examine individually, when needed. Within the instructional component, context links may send to description of individual items, whenever they are mentioned and also according to the presumed user knowledge of each of them. We will focus, from now on, on the instructional Section of the manual.

3. Outline of GeNet

The architecture of GeNet is shown in figure 1.

Figure 1: Architecture of GeNet

The generation process starts with the definition of a communicative goal. Starting from this goal, a discourse planner performs the task of information selection and basic organization, according to the user information needs that are represented in a stereotype-based User Model (UM). A hypermedia generation algorithm transforms the resulting discourse plan into an hypermedia structure. DB, KB and UM Servers store and provide the following information:

i) domain knowledge: data about the descriptive and instructional parts of the manual, in textual ot pictorial form;
ii) plan operators and hypermedia transformation rules;
iii) stereotypes, user profiles and adaptation rules.

The last component provides various types of data about the user to the two generation components: planning takes into account the user knowledge in the domain and the task to be performed; hypermedia surface generation considers (as we will see later on) other features, such as the level of experience and the hypothesized preferences. These data are acquired explicitely from the user, through a 'user presentation' page which is displayed at the beginning of interaction. Let us now describe in more detail the system components and the methods employed.

3.1. The discourse plan

Our planning algorithm is described in detail elsewhere (De Carolis et al, 1996); compared with other discourse planning methods, it has the following distinguishing features:

- plan operators represent, at the same time, how a communication goal (mentioned in the Header) can be decomposed into subgoals (in the Decomposition) and which Rhetorical Relation links the discourse segments that originate from these subgoals. The subgoal which produces the RR's nucleus is denoted with an 'N'; the RR may have several satellites, all denoted with an 'S': for instance, an 'Elaboration Process Step' with a nucleus and several satellite steps is represented in a unique plan operator;
- planning is made hierarchically according to a double adaptation: to the Speaker and to the Hearer. Data about these users are stored in two user models. A meta-level rule-based planning decides whether to further expand the plan obtained at each abstraction level, by also considering complexity of the plan generated so far. The complexity of a node in the plan tree is a function of depth and breadth of the subtree of which the node is the root, and of the type of leaves (text vs images or other).

The result of the planning step is a discourse plan (that we will call Dplan), which is represented as a tree with a set of data associated with nodes.

Def:

Dplan : $<$ T, Go, Db, Fo, Rr, Cc, Ro, F_d $>$, where:

- T is an *ordered tree* whose nodes are denoted with N = $\{n_1, n_2, ..., n_m\}$
- Go is a set of *communicative goals*;
- Db is a set of *data* (phrase fragments or images) retrieved from a database;

- Fo is a set of *concepts or device elements* that correspond to discourse focuses;
- Rr is a set of *rhetorical relations RR* {Sequence, Joint, Contrast, ...}
- Cc is the set of *complexity classes* of a subtree {L, M, H}, which are defined as a function of its breadth and depth (L = 'low', M = 'medium' and H ='high');
- Ro is the set of *roles* that a node may have in a RR: {N, S}, (N = 'nucleus', S = 'satellite')
- *Fd*: $N \rightarrow Go \times (Db \cup \{\varepsilon\}) \times Fo \times Rr \times Cc \times (Ro \cup \{\varepsilon\})$ is the *node description function* which associates, with each node of T
  - *the communicative goal* of the plan operator which generated the node,
  - *a phrase or an image* (if n_i is a leaf-node)
  - *the discourse focus*
  - *the rhetorical relation* among n_i' s children,
  - *the complexity class* of the subtree of which n_i is the root,
  - *the role* of n_i in the rhetorical relation associated with its father-node (N or S) if $n_i \neq n_1$.

Figure 2: a portion of the discourse plan for the blood pressure measurer manual.

Figure 2 indicates part of the plan for the instructional section of a blood pressure measurer manual. This subtree illustrates, in particular, how to "put the armband and position the arm" (this is the communicative goal of root node n_1). This goal is achieved by decomposing the discourse into four segments (n_2 to n_5), which are linked by a RR of *ElaborationProcessStep* and whose goals are, respectively:

Request (put the armband and position the arm), in n_2
DescribeAction (prepare the cuff) in n_3
DescribeAction (wrap the cuff around the arm), in n_4
DescribeAction (position the arm correctly), in n_5.

The discourse focus is 'cuff' for node n_1 and n_3, 'armband' for n_2, and so on. The complexity of the subtree from n_2 and n_3 is=Low; the complexity of n_4, n_5 is High. All nucleus are marked with a circled dot, in the figure. The RR of *ElaborationProcessStep* relates a nucleus describing the process (n_2) to satellites (n_3, n_4, n_5) that represent the steps necessary to perform the process itself. Other RRs represented in this tree are *ElaborationObjectAttribute, Motivation, Contrast* and *CauseResult.* In this example, a *Background* is used to relate a text to an image that illustrates its content. As we mentioned before, our RRs can have several satellites.

3.2. Hypermedia generation criteria

Our hypermedia (that we will call Hypm) can be represented as a graph structure: pages are associated to nodes and arcs represent links between couples of nodes, in the form of anchors, buttons or icons.

Def

Hypm: < P, Ti, He, Bo, Ft, Bu, Hw, *Co, Re* >

- P is a finite, non empty set of *pages*: $\{p_0, p_1, ...\}$; p_0 denotes the 'home page';
- Ti is a set of *titles;*
- He is a set of *heads;*
- Bo is a set of *bodies*;
- Ft is a set of *footers;*
- *Co* is a *page-content definition function:*
  Co: $P \rightarrow Ti \times He \times Bo \times Ft$ which associates to each page a Title, a Head, a Body and a Footer;
- *Re*: $P \rightarrow 2^P$ is a *link definition function* which associates with each page a set of related pages.

Hypm is built by an algorithm which exploits Dplan knowledge to create Hypm pages, and links among them. This algorithm performs the following functions: (i) establishing the number of Dplan nodes to be aggregated into a Hypm page, (ii) transferring into Title, Head and Body of each page information items associated with Dplan nodes; (iii) inserting anchors in the Body and Buttons in the Footer, with their URLs. In this process, Rhetorical Relations (RRs) associated with Dplan nodes play a crucial role: they establish how to structure the Body, how to insert anchors and how to compose the Footer according to the selected navigation path. This process is driven by heuristic criteria that we derived from guidelines for interface design and from Web construction manuals (Thimbleby,1991): these criteria are embedded in the generation algorithm.

a. page complexity control*:*

Information shown in each page is the result of a compromise between two contrasting criteria: a page should contain the maximum amount of 'autoconsistent' information, so as to avoid fragmenting knowledge in Dplan into too many Hypm pages and should be displayed, if possible, in a single layout, so that it doesn't overwhelm the user. We call 'autoconsistent' the information associated with the components of a RR (nucleus and satellite): a page p_h then describes, in a synthetic or detailed way, nucleus and satellites of the RR which is associated to the node n_i from which p_h is derived. For example: a page which illustrates a discourse segment guided by a *ElaborationProcessStep* describes first the overall process; the steps necessary to perform it are described in the same page, either synthetically (if the corresponding node's complexity is 'High') or in a detailed way (if it is 'Low): synthetic descriptions are developed in subsequent pages. Given a node n_i in Dplan and the subtree starting from n_i, the number of Dplan nodes to be aggregated into the p_h page of Hypm (only n_i and its children, or also its descendants) depends on the complexity associated to the children nodes.

b. consistency control:

Consistency is insured by uniform page design: pages corresponding to 'similar' discourse segments should have a 'similar' layout; buttons and icons representing similar links in different pages should be placed in the same position and rendered with the same image or text; navigation paths among pages corresponding to similar discourse segments should be similar; same hotwords shown in different contexts should send to the same page. The similarity criterion we adopt in GeNet is based, again, on the RR which guides the discourse segment: pages originated from the same RR have a similar layout: same components and same space arrangement and presentation for each of them. The structure of links among pages is defined, as well, from RRs among discourse segments: segments guided by the same RR are described in sets of pages with similar navigation paths, suggested by similar buttons or icons. For example: 'Next item' buttons always refer to the next item of a *Sequence*; a 'Next step' to the next step of an *ElaborationProcessStep*, and so on.

c. navigation styles

As we mentioned in the Introduction, the effectiveness of the navigation style is related to the user characteristics and to the task they have to perform. We included in GeNet some of the navigation modes that cognitive psychologists claim to be the most efficient for learning tasks: a '*guided tour*', in which the Speaker's intentions are described clearly and the sequence of page presentation is rigid and a '*free*' mode, in which users are left free to select the parts of the hypermedia they want to examine; in both cases, the Speaker's intention can be either described initially and in synthesis, in an '*index*' page, or, in more detail, in a hierarchy of intention-description pages. Users should also receive a level of orientation, in their navigation, which depends on their experience in the domain and in hypertexts in general. Orientation supports may be provided in several forms; by 'connection pages', that summarize what the user has seen until that point and anticipate what he is going to see next, to describe passages from a comunicative goal to the other; and, by labels associated with buttons and icons and initial goal mapping.

4. The generation algorithm

The hypermedia generation algorithm employes a set of micro and macro transformation functions.
At a **micro** level, the following elementary functions are applied:

- λ (x) is an ATN-based *linguistic transformation* that builds up a phrase, a clause or a sentence from a word, a phrase or a discourse segment or a list of words x;
- π (y) is a p*ictorial transformation* of an image; for example: zooming on a part of it, reducing its size, changing its colours;
- μ (w_1, w_2, ..., $w_{i)}$ is a *multimedia combination* of sentences, images or other multimedia items; for example: attaching a text label on the bottom of an image;
- α (t,p_j) selects a string or a substring t as a link anchor and associates it to a page p_j;
- β (v,$p_{j)}$ selects a button v among those available, and links it to a page p_j.

At a **macro** level, functions are employed to generate the page elements, thus defining the page-content definition function *Co(p_i)*. Creating a page requires deciding, first, how Dplan nodes are to be aggregated into Hypm nodes; this process is influenced by the RR, the complexity of the information to be provided to the user for achieving the communicative goal and the selected navigation mode.

Our micro-level generation functions are, still, rather rudimentary, as we concentrated, in this phase of research, on macro-level functions (overall page organization and navigation styles); linguistic transformations are just obtained by just immersing the string associated to the 'Communication Goal' topic or to the 'Data' object-slots into a linguistic template, that is selected according to the Goal purpose and to the page Section. Multimedia combination arranges texts and pictures into a space relationship that depends on the RR. For instance: in a *Background* RR, the text is placed at the leftside of the image that illustrates it; in an *ElaborationProcessStep*, step descriptions are arranged sequentially in a numbered list,...and so on.

The following macro functions are applied to create a page:

- `Title`(n_i,p_h): the title identifies the page and orients the users in the navigation history; this function composes the title of page p_h by applying a linguistic transformation to the Goal associated to the plan node n_i.
- `Head`(n_i,p_h): the header cues the user to the page purpose; this function composes the header of page p_h by combining linguistic transformations of the Goals associated to the plan node n_i and to its father (if n_i is not the plan root).
- `Body`(n_i,p_h): this is the most complex section of the page, as it is in this section that the discourse tree nodes are aggregated and structured. The discourse tree is visited in pre-order, and nodes are aggregated according to: i) the RR associated to the root of the examined subtree, ii) the complexity of its children and iii) the selected navigation mode.

- in the *indexed* mode, a home page is generated from n_1 and all its children of high complexity. The body of this page shows a list of items which are developed in child pages; navigation in these pages can be 'linear' or 'hierarchical';
- in the *free* and the *guided* modes, the intentional structure of the manual is mirrored in a hierarchy of pages. In the free mode, links to lower-level pages is produced by a set of anchors that are produced by applying the alpha micro-function, as well as links to the descriptive part of the manual. In the guided mode, buttons send to the next page to visit.

• Foot(p_h, n_i, p_j): the footer allows creating links in the *guided* navigation mode; this function selects a set of buttons according to the RR and associates to each button a link to a page p_j, by the micro function β (v, p_j). If, as in the previous example, the RR is an ElaborationProcessStep, the footer buttons indicate the possibility for the user to go to the "Next Step" or to the "Previous Step", while in the case of a Sequence this link will be indicated as "Next Item" or "Previous Item". This can be viewed as a form of cognitive orientation support for the user who is navigating in the hypermedia; the level of support can be strenghtened by mentioning, in the button label, the name of next step or item.

We show, in Figure 3, an example of generation of a page from the subtree starting from node n_1 in Figure 2, in the *guided* and the *free navigation modes.*

mettere solo una pagina?

Figure 3: An example of page layout

The main RR is an Elaboration Process Step: the nucleus describes the process and the satellites the individual steps. The page Title is obtained by a linguistic transformation function that inserts the Goal topic of the tree root, which is out of Figure 1 (item1: 'use the blood pressure measurer correctly') and the Goal topic of n_1 (item2: 'put the armband on and position the arm') into the simple linguistic template 'How to <item1> : <item2>'. The Body is obtained by a multimedia transformation that combines a linguistic transformation of the process (data associated to object n_2) with an ordered list of step descriptions. The first step description is obtained, in its turn, by a multimedia combination of items associated to children of n_3 (whose complexity is 'Low'); descriptions of steps 2 and 3 are only synthetical, because nodes n_4 and n_5 are of 'High' complexity. In the *free navigation mode* (rightside), an anchor is inserted in the descriptions of these steps, to enable the user to have further details if needed. In the *guided mode* (leftside), a button sends the user to the page that describes the second step: the label of this button orients the user, by anticipating the content of that page.

4. User-adaptation

In hypermedia, dynamic adaptation can be provided at two levels:

- at the *presentation* level, by customizing the node content and layout (text, media allocation, information selection, etc.);
- at the *navigation* level, by customizing the navigational support and paths (link hiding, reordering of links, adapted guided tours, etc.). Unexperienced users are guided in examining the manual systematically, so as that they learn all aspects of instruction and understand the purpose of any piece of information; experienced users are left more free in examining the manual so that they concentrate only on those aspects about which they are not familiar. Users also receive a level of orientation, in their navigation, which depends on their experience in the domain and in hypertexts in general.

Although the two levels of adaptation are conceptually different, in GeNet, they are strictly related. Adapting at both levels is done by exploiting knowledge about the user's preferences in the user model; our user model is a simple stereotype that is triggered by age, educational level and experience in the application domain and hypermedia. As micro and macro generation functions are parametrized, user model settings are employed to set up the following Hypm features:

- details to be included in a page (how many Dplan nodes to aggregate in a Hypm page);
- page layout (how to generate Title, Header and Body). For example, linguistics transformations of information items can be performed by employing terms with which the user is familiar;
- anchor and button links;
- level of orientation support;
- the last two features correspond to the navigation path and mode. For example, for inexperienced users at their first interaction with the system, a "guided mode with connection pages" is proposed;
- checking of whether the user skipped information items that are supposed to be relevant in the instruction process.

Figure 4 shows an example of navigation mode for the Dplan in figure 2: the user is aged, with low domain and hypermedia experience, and desires seeing figures. A guided navigation mode with strong orientation, medium fonts and high level of detail is then selected by adaptation rules.

Figure 4: an example of adaptation in the navigation style

After the first page (that describes the process 'put the armband on and position the arm'), the user is guided to see pages corresponding to steps 1 (in detail) and 2 (in synthesis). Details of step 2 are described in the three subsequent pages (2.1: 'position the armband ', 2.2: 'position the garment' and 2.3: 'fix securely the armband'). Description of step 2 is now finished and description of step 3 ('position the arm correctly') can begin: the user is oriented about the phase of instruction to which hypermedia has come, by a 'summary page' which synthetizes the goals of the completed step and of the next one.

5. Conclusions.

Generation of user-adapted hypermedia can go beyond crude filtering of raw Web material if an appropriate knowledge base and refined multimedia generation methods are employed (Inder and Oberlander, 1994; Brusilovsky et al, 1996). In the generation of concept descriptions, this KB can be a hierarchy of concepts or a semantic network of items to be described. We proved that, in the case of instruction manuals, knowledge contained in a RST-based discourse plan and in the user model can be employed effectively to build hypermedia by ensuring internal consistency of pages and global consistency of the entire document. Our algorithm enables producing several navigation paths and page presentations from a unique discourse plan; the advantage is that, after text planning, one may decide whether to generate a one-shot document, a hypermedia or both. One can also generate several hypermedia from the same plan at the same time, each tailored to a specific user category.

Our future work aims at:

- studying how to improve dynamic presentation of information items by refining media allocation and integration methods;

- verifying, through a set of evaluation studies in cooperation with experimental psychologists, which combination of media, which navigation modes are the most effective in reaching the desired goal, according to the user characteristics and preferences. The results of these experiments will be employed to verify whether the three modes and the page presentation that we generate with the present algorithm are effective from the instructional viewpoint, and how the user acceptance of the generated hypermedia can be achieved best.

References

P Brusilowsky, E Schwarz and G Weber, A Tool for developing adaptive Electronic Textbooks on WWW. *WebNet96 Proceedings*:

B De Carolis, F de Rosis, F Grasso, A Rossiello, D Berry and T Gillie: Generating recipient-centered explanations about drug prescription. *Artificial Intelligence in Medicine*, 8, 1996.

J Eklund and R Zeiliger: Navigating the Web, possibilities and practicalities of adaptive navigational support.

E H Hovy: Automated Discourse Generation using discourse relations. *Artificial Intelligence*, 63,1993.

R Inder and J Oberlander, Applying discourse theory to aid hypertext navigation. *ECHT94.*

W C Mann and S A Thompson: Rhetorical Structure Theory, towards a functional theory of text organization. *Text*, 8, 3, 1988.

M Milosavljevic e R Dale: Text Generation and User Modeling on the Web.

V Mittal and J Moore: Dynamic generation of follow up question menus: facilitating interactive natural language dialogues. *Proceedings of CHI95*

H Thimbleby: User Interface Design. *ACM Press,Frontier Series,* 1991