Sources of Flexibility in Dynamic Hypertext Generation

Alistair Knott † Chris Mellish ‡ Jon Oberlander † Mick O'Donnell ‡
† Human Communication Research Centre, 2 Buccleuch Place
‡ Department of Artificial Intelligence, 80 South Bridge
University of Edinburgh, Scotland
{A.Knott | C.Mellish | J.Oberlander | M.ODonnell}@ed.ac.uk
Fax: (+44) 131 650 4587

Abstract

'Dynamic hypertext' is hypertext which is automatically generated at the point of need. A number of NLG systems have now been developed to operate within a hypertext environment; and now that these systems are becoming widely available on the World Wide Web, it is useful to take stock of how well-equipped NLG technology is to work in this new domain. A generation system in a hypertext environment faces a specific set of requirements; here, we discuss those requirements, and the resources that can be provided to help meet them. Examples are drawn from a number of systems, including our own prototype, ILEX-0. We conclude by indicating that the major benefit of such systems could be in the way that they combine flexibility with the *illusion* of user control.

Keywords: content selection, text planning, applications, hypertext

1 NLG in a hypertext environment

1.1 Static vs Dynamic Hypertext

Within the last ten years there has been a upsurge of interest in hypertext as a medium for on-line access to written documents. As a central objective of Natural Language Generation (NLG) is on-line *creation* of written documents, it is hardly surprising that a number of generation systems have been developed which use hypertext as their interface.

Conventional hypertext is **static**, in the sense that—even though each reader may sample different parts of it—once a hyperdocument has been constructed by its authors, its content and form do not change. However, applying NLG to hypertext allows the possibility of **dynamic hypertext**, where the document is created as demanded by the user. This approach allows the hypertext pages seen by the user to be customised in relation to the browsing context.

In some cases, customisation is only in relation to content—the page brings together information from various on-line sources, but presents it in a single standard way. Other systems, using more advanced NLG techniques, allow the text to be customised in terms of both content and presentation form, being sensitive to such factors as the **user model** (characteristics of the user), the **discourse history** (a record of information presented so far), and the **system's goals** (what the system means to achieve).

Amongst the latter systems, the degree of flexibility supported by the text generation component varies considerably. In some systems (such as StrathTutor [Kibby and Mayes 1988]) and Trellis [Stotts and Furuta 1989]), the hypertext pages are hand-crafted in advance, but the way they are linked together varies from session to session depending on the interests of the user. In other systems, the hypertext pages themselves are generated at runtime, as well as the links between them. In many current systems, a mixture of these two methods is employed.

Clearly, there are many different ways of introducing flexibility into a hypertext system. Our concern in this paper is to outline the most important of these, indicating which have already been explored, and which others deserve closer attention within the NLG community.

1.2 Existing Dynamic Hypertext Systems

A number of hypertext systems incorporating NLG technology have already been developed. In this paper, we will make reference to the five systems listed below.

ALFRESCO [Carenini et al. 1993]: provides an interface to a videodisc of Italian frescoes and monuments. It supports natural language input and output, and generated text contains hypertext links into an existing static hypertext.

IDAS [Reiter at al. 1995, Reiter et al. 1992]: delivers simple technical documentation online via a hypertext interface.

MIGRAINE [Buchanan et al. 1995]: generates information and reactive explanations for a patient, based on their medical record.

PEBA-II [Dale and Milosavljevic 1996, Milosavljevic et al. 1996]: interfaces with a zoological database. Delivers either a description of a single user-selected animal, or a comparison between two user-selected animals.

PIGLIT [Binsted et al. 1995]: allows a patient to interface to their medical records, and receive personalised explanation.

In addition to these systems, we will introduce our own prototype system ILEX-0, which is being developed specifically to examine issues arising in the domain of dynamic hypertext. The system is designed to simulate the interaction of a museum tour guide and a browsing visitor; this interaction is described in more detail in Section 2. In Section 3, the browsing task provides the context for a general discussion of the factors which influence the ability of a hypertext system to function dynamically, and in the light of this discussion, some observations about existing dynamic hypertext systems. Section 4 then draws some conclusions concerning the real potential of dynamic hypertext.

2 A sample domain: Museum guided tours

A central attraction of the hypertext interface is its suitability for the task of information **browsing**. By following hyperlinks, the information seeker (i.e. the user) can easily navigate to new areas of interest. At the same time, for the reasons already mentioned, there are advantages in being able to provide the information requested by the user in a dynamic fashion. If we look for an analogy in the real world to the kind of interface we are aiming for, an interesting candidate is the dialogue between a museum tour guide and a browsing visitor. The ILEX (Intelligent Labelling Explorer) project has been set up to study this domain, and to reproduce some of its distinctive features in a hypertext system. To achieve this goal, our system must:

Support mixed initiative dialogue There is a degree of 'dialogic' interaction between the tour-guide and the visitor. While there are quite restricted directions that the conversation can take, each party can take the initiative in pursuing a direction.

Be interesting The visitor is free to ask about whichever objects take his fancy; if an object does not hold his interest, he will move on to another, or finish.

Be informative The system has all the information, and has the goal of getting across some basic messages about the content of the gallery.

Clearly, the system needs to alter its descriptions in the light of the objects the visitor has already heard about, and to have some idea about what the visitor might find interesting. This is an increase in functionality compared to existing static museum hypertexts, for instance as created in the UK by the Hunterian Museum and the National Gallery.

Our point of departure has been to gather transcripts of real guided tours, in collaboration with the National Museums of Scotland. We ran two 'Curator-of-Oz' interviews, where a curator of the modern jewellery gallery was

V:	Can we look at Case number 7 now, object number 10, this one here...
C:	Yes, you've made a link with the first piece that we looked at, which is the idea of a piece of jewellery which is also a work of art and a sculpture...[describes jewellery].
V:	Can you say more about that object?
C:	Yea. This one was made by Roger Morris. [Describes jeweller].
V:	So is there any other object which is particularly relevant to this one?
C:	Yes...[points out and describes two other brooches]. It was work like this which directly inspired work like the Roger Morris brooch we looked at earlier.

Figure 1: Part of the 'Curator-of-Oz' transcripts

Item 342040 : a brooch on pyramid mount

This is a brooch with mount made of acrylic, 18 carat gold and silver. It was designed by Roger Morris (London) in 1976.

With a piece like this, the boundary between 'jewellery' and 'sculpture' or 'art' starts to become quite indistinct. One important theme across 20th Century jewellery has been what to do with a jewel when it is not being worn. From the 1970s onwards, jewellers have started exploring the idea of turning jewellery into sculpture -- so that you can hang it on the wall, or prop it on the mantelpiece when you are not wearing it. This piece works equally well whether it is being worn or being displayed (as at present). To see other examples of objects which are both jewellery and art, click on the following: 343302 990811 342743 237307 .

What you are looking at is not all a piece of jewellery. Part of it is actually its mount, so that you can hang it on the wall, or prop it up on your mantelpiece when you are not wearing it. To see other examples of objects which are mounted, click on the following: 343302 990811 342743 237307 .

Roger Morris is well known as a teacher of Jewellery Design. He has been quite influential on the field as a whole.

Figure 2: A Sample Hyper-Tour

System	Representation Granularity	Linguistic Sophistication	Discourse History	User Model	Goals	Many Texts	Many Options
ALFRESCO	KR+Canned	Full+Canned	Initially	Yes	User	Initially	No
IDAS	KR+Intermed.	Full+Template	Locally	Yes	User	Yes	No
PEBA-II	KR+Canned	Template	No	Yes	User	Yes	Yes
PIGLIT	KR+Canned	Template	No?	Yes	System	No	No
MIGRAINE	KR	Template	Yes	Yes	Mixed	Yes	Yes
ILEX-0	KR+Annot	Template	Some	Yes	Mixed	No	Yes

Figure 3: Dimensions of Flexibility in Existing Systems - Representation Granularity and Linguistic Sophistication (Section 3.1), Sensitivity to Discourse History (Section 3.2), User Modelling (Section 3.3), Location of Goals (Section 3.5), Many Texts and Many Options (Section 3.4)

asked to talk about a series of objects in response to simple prompts from a visitor. Figure 1 shows an extract from one such interview. V is the visitor, and C is the curator. Based on the information extracted from these tours, and additional database material, a prototype system ILEX-0 generates hypertext pages; for a sample page, see Figure 2.

ILEX-0 is still far from generating discourse of the sophistication shown by a real museum curator. But it should be clear that any hypertext system designed to simulate such discourse must make use of a dynamic generation component of some sort. In the following section, we consider the kinds of flexibility which this component should ideally provide.

3 Parameters of flexibility in hypertext generation

We indicated in Section 1.1 that an NLG system can exploit a number of techniques to generate hypertext. In this section we will attempt to isolate a number of parameters which affect the flexibility of the resulting dynamic hypertext. Although obviously such a short table cannot do justice to the detailed contributions of these systems, Figure 3 summarises our understanding of how the above six systems vary in terms of the parameters, which are discussed in the following sections. In each case, we assume that the parameter influences the quality of the text, and we ask whether NLG needs to investigate the parameter in greater detail in order to produce good hypertext applications.

3.1 Granularity of information representation

A fundamental factor influencing generator flexibility is the form of representation of the information to be expressed. At one extreme, the information can actually be *stored* as text, in which case there is no scope at all for flexibility of expression; this is essentially the scenario for 'canned' text. Scope for contextual variability is introduced by establishing an *independent* format for information storage, and a flexible mapping algorithm from this to surface text. The linguistic sophistication required of the system obviously depends on the level of representation chosen.

In fact, the 'degree of cannedness' of a system's output is on a sliding scale. A static hypertext is made up entirely of canned text. In a slightly more flexible system, individual sentences or paragraphs may be canned, while their organisation and inclusion is constrained by some theory of discourse. (For instance, a system could store canned sentences organised in terms of an RST tree; the generator could then use the RST information to select and linearly organise the canned segments.) More flexible still, a system could use canned text only at the phrasal level, for example in a template-based sentence generation system, or in a phrasal lexicon. Finally, a system could dispense with canned text altogether, and use full NLG technology. ALFRESCO uses general generation methods for the first screen produced, but canned representations for the other nodes accessible by hyperlinks. Most of the systems discussed in this paper use canned text in some places though also being able to generate variable text. IDAS uses a number of intermediate forms that mix canned text with standard knowledge representations.

Concern with the appropriate degree of cannedness for NLG systems is of course not limited to hypertext generation. It is currently the focus for much debate in all areas of NLG, especially where the development of practical applications is an issue (see e.g. [Reiter et al. 1995]). But as the dimensions of flexibility we are interested in depend on this issue, it is important to discuss it.

ILEX-0 represents information at two distinct levels on the cannedness scale. For fine-grained detail, the system uses a database made available by the museum, containing entries for the 120 or so jewellery items in the Modern Jewellery exhibit. These entries have been parsed in and represented in a KB-accessible form, as shown in Figure 4. The generation of sentences based on this information is currently effected via a simple template-based mechanism.

Secondly, the transcriptions of the curator tours have been hand-processed to extract canned **stories** about individual jewellery items, the designers of the items, and

```
(def-jewel
  :case '7-10
  :isn 342040
  :code 'a.1987.284
  :short-title "a brooch on pyramid mount"
  :date "1976"
  :class 'brooch
  :class-text "brooch with mount"
  :materials '(acrylic yellow-material
              white-metal)
  :designer 'morris01
  :dimensions "5.20\" L"
  :properties '(mounted)
  :place "ENGLAND: London")
```

Figure 4: A Sample Database Entry

```
(def-story
  :id 'PAPER-1
  :trigger 'PAPER
  :type :PROPERTY
  :message 'MANY-KINDS-OF-JEWELLERY
  :singular-label "item made of paper"
  :plural-label "items made of paper"
  :text "Jewellery need not be made of
        metal. This item, quite unusually,
        is made from paper. This shows
        that a wide range of materials
        can be used and still be called
        'jewellery'.")
```

Figure 5: A Sample Story using only Canned Text

more often, classes of jewellery (e.g. jewellery made in a particular style). Figure 5 shows one example.

However, given the limitations of current NLG systems, if it is feasible canning offers the most reliable way of ensuring high-quality text output. And, since it is much easier to enter information in the form of text than using a complicated knowledge representation formalism, canning also facilitates the authoring task. But of course, canning fails to support genuine flexibility. In the ILEX-0 system, we have sought to allay this problem by including in the canned text various types of **annotation** which allow a certain amount of local customisation. Below are two possible annotations, only the first of which is presently included in ILEX-0:

- **Conditionalisation**: Parts of the text can be conditionalised, indicating that the text is to appear only if the condition is met, e.g.

  ```
  "They used the very best designers
  to design jewels for them.
  <cond Designer=KING01> Jessie King
  was one of their best designers.
  </cond>"
  ```

  This allows canned text to include comments about particular entities if they are relevant at that point in the discourse.

- **Reference**: Entities which are defined in the knowledge-base can be replaced with an annotation, which allows the generator to choose a referential expression appropriate for the context of generation, e.g,

  ```
  "#(:ref KING01) was one of their
  best designers".
  ```

Further types of annotation will be added as the ILEX system develops. We plan to explore whether, for some applications, the medium of 'annotated text' can be a useful compromise between canned text and full NLG.

3.2 Sensitivity to Discourse History

At any point in the dialogue, a dynamic hypertext system should be able to take into account the discourse which has already taken place. This can be exploited in several places; we here discuss two.

3.2.1 Referring Expressions

The readability of a text is partially determined by the appropriate use of referring expressions.The continual use of an entity's full name produces bad text, just as does the use of a pronoun without any preceding referent. Since dynamic hypertext is being produced chunk-by-chunk, the system needs to keep track of prior reference forms for an entity, and choose a form for the current mention which is appropriate. For instance, if we have recently mentioned an entity of type T, and we now need to introduce an object of the same type, then we could refer to it as *another* T. There are already several NLG systems which incorporate sophisticated algorithms for the generation of referring expressions such as these; see for instance [Dale 1992] and [Horacek 1995]. PEBA-II and PIGLIT have no discourse modelling, whereas ALFRESCO uses it within the first generated node only. IDAS tracks the discourse history, but only within the current hypertext node being generated. We need a system which can keep track of reference across nodes.

In particular domains, we might expect other kinds of sensitivity to preceding context. For instance, if the hypertext session is intended to simulate a spatio-temporal domain, then it makes sense to talk about 'returning to' locations which have already been visited. But if the tour is of a more conceptual space, this is not necessarily what is needed. One of the domains in IDAS is *bicycle mechanics*; it might often be inappropriate to talk about *returning to* the wheel rim of a bicycle.

3.2.2 Content Selection

It is also desirable that the *content* of the text produced is sensitive to what has been said in the preceding discourse. We assume that the system aims to tell the seeker something interesting—but what is interesting depends partly on what has already been said. Some things become interesting only because of prior context. For instance, the system might have a very informative comparison to make between two classes of objects in the domain being browsed; but if the seeker has not already come across instances of each class of objects, the comparison might not be interesting or even intelligible. The system's decision about whether to include the comparison might thus be strongly influenced by prior context.[1]

On the other hand, some things cease to be interesting if they have already been discussed, due to repetitivity. This problem, called here the **redundancy** problem, affects static hypertext systems: since a series of hypertext pages can be browsed in a variety of sequences, a static system cannot know at any point what the user has already been told. Information essential to the interpretation of the present text may need to be repeated at each point it is needed. As a result, the reader may become bewildered by the mass of information presented, unable quickly to determine what is new and what is a restatement of what he has already read. A dynamic system however can keep track of what has been told to the user, and not repeat such information (except, naturally, for rhetorical reasons).

Of course, when we decide not to re-express old information, we can still provide a backward-pointing link. If a story has already been given (an earlier piece of jewellery had the same property or designer), then repeating it verbatim can have a serious effect of the user's perception of text quality. Because of this, the second time a story is triggered in ILEX-0, the story is not presented, but rather, a pointer is given back to the prior piece where the description occurred. For instance:

> Like the earlier piece of jewellery (240384), this piece is also a piece made for Liberty & Co.[2]

A third content-suppressing mechanism implemented in ILEX-0 is for handling terminology definition. The system keeps a list of terms defined for the current user, and avoids re-defining these terms.

The question arises as to whether hypertext applications demand more work on the discourse history parameter—supporting referring expressions and content selection—than has hitherto arisen in NLG. It might, for instance, be necessary to model the importance of what is physically visible on the screen or the fact that a text may not have been completely read before a new one is selected. The tentative answer—arising particularly from the need dynamically to suppress content—is that this is an area where further work is required.

[1] An alternative approach to the generation of comparisons is provided in PEBA-II. In ILEX, comparisons are to be produced spontaneously where the system judges they are appropriate; while in PEBA-II, they must be asked for by the user. The two systems differ in terms of which participant is in *control* of decisions about the information provided; for more discussion on this subject see Sections 3.4 and 3.5.

[2] Here, some content is successfully replaced with a backward pointer; however, repetition is still rampant because of the primitive templates (see Section 3.1).

3.3 The User Model

A useful kind of flexibility in any NLG system involves keeping a model of the reader of the text. Texts produced for an expert will be very different from those produced for a novice, even if they are on the same subject. Likewise, texts for an adult audience will differ from those intended for children. In the domain of guided tours, user models will clearly be important; in fact, the National Museums of Scotland have a ramified visitor classification scheme, with four categories of visit (including *tourist*, *formal educational*), fifteen types of visitor (including *9-year-old*, *tourist with Scottish connections*), and four levels of visitor interest (including *baby browser* and *specialist*). Many NLG systems have taken such dimensions of variation into account: see for instance [Paris 1993].

An important dimension of variability in hypertext generation relates to *control* over the user model. Changes to the user model can be under the control of the system or the user. We might envisage a form-based **preferences page** where the user can configure his own user model, either before browsing begins, or at any point within browsing. The ILEX-0 prototype allows the user to specify the degree of interest he has in a set of themes. These preferences are then used to drive content-selection. We also plan to experiment with user-changeable parameters, such as 'longer/shorter texts', 'more/fewer comparisons' etc. These choices are roughly analogous to meta-comments in ordinary discourse.

Alternatively, updates to the user model can be system-initiated, and driven by observations about the options chosen by the user during browsing. If the user often selects the *tell me more about this* button, then the quantity of text given at each point can be increased. A user who often asks for terms to be defined can be offered more definitions within the text. A user model is thus built up by behavioural observation. This approach can also capture changes in the user's interest as they occur.

ALFRESCO maintains an interest model for its user which can affect what is generated in the first textual response. However, a system of this kind which sets up a network of hyperlinks in its entirety in response to a single query from the user, will not be able to take full advantage of a dynamic user-model. The ideal may be a system where pages and links are created one by one, and the user has the opportunity for input at every stage.

The theory of user modelling in NLG is still at a relatively early stage of development. However, state-of-the-art techniques [**?**] should be sophisticated enough to deal with many of the requirements of dynamic hypertext mentioned above.

3.4 User Freedom

In a hypertext system, the user's requests for follow-up information are mainly in the form of 'mouse clicks' to new hyperlinks. From the user's point of view, there is a degree of flexibility in the system even if the hypertext is fully generated in advance, because he has a choice about what to look at next. User freedom in conventional hypertext is somewhat of a double-edged sword: giving users the choice about where to navigate is helpful, but it can also leave them uncertain about whether they have picked up all the important information on offer. However, in a dynamic hypertext, user freedom is not just to do with the range of follow-up hyperlinks. A second, less visible, kind of freedom results from a system's ability to decide at any point between a large range of candidate texts for generation.

In some systems, all the possible texts have been anticipated explicitly in advance, even though the detailed contents may be dynamically generated. An example might be a system that offers weather forecasts for a fixed set of regions but where the individual forecasts may change to reflect up-to-the-minute information. The PIGLIT system is rather like this, although in the medical domain. More flexibility is offered if the text produced depends on a set of independent dimensions. In IDAS, for instance, there were over 5000 potentially different texts that could be generated (ignoring some dimensions of variation). PEBA-II has similar flexibility, in that its comparison texts depend on *two* animals that are chosen by the user independently. ILEX-0 is nearer to the PIGLIT model, because its primary mode of access simulates the actual layout of a museum gallery, where the user has the simple choice of which of a set of objects to look at next. Of course, as soon as user modelling, sensitivity to discourse history, etc., are reflected in the texts, then the number of possible texts grows quickly (indeed, the above IDAS figure takes into account some user modelling dimensions). This potential freedom may not be obvious or available to the user, however.

Just as a speaker must allow a hearer opportunity to respond, it is important for the system to generate a fair selection of new links to explore, to ensure that the reader finds something to his liking. In ILEX-0, the user is able to pick any object in the gallery to look at next, which is intended to allow the same kind of random movement that a user might make in a real museum gallery. On the other hand, it might also be important not to generate too many links to follow up, as this might make navigation confusing. In IDAS, for instance, an explicit decision was made to limit users to following certain 'hyperschemas', rather than giving them the full choice. Where the set of possible texts has not been anticipated in advance, a more complex (perhaps menu-based) query system to allow the user to choose a text may be provided (as in PEBA-II).

Thus, the ability to specify a restricted subset of potentially very large set of user follow-up actions is a particular feature of dynamic hypertext. It follows that the whole issue of 'user freedom', and ways of anticipating user actions will require further work in NLG. Users' freedom—or their perception of freedom—is intimately bound up with the ways in which systems realise their own goals, and it is to this that we now turn.

3.5 The System's Goals

For a hypertext system to be successful, it needs not only to succeed from the user's point of view (to be interesting), but also to succeed from the provider's point of view (to be informative).

In general, the provider has her own agenda in creating the hypertext, for instance to persuade the user to buy something, or to educate him in a particular way (this cannot be seen obviously in query systems like PEBA-II, IDAS or ALFRESCO, but is apparent in a system with more of a tutorial role, as in MIGRAINE and PIGLIT). In a browsing situation, however, the provider cannot plan ahead towards a complete text which satisfies her goals, since the reader's actions cannot be predicted in advance. Goals must thus be pursued **opportunistically**—the system achieves aspects of its goals as the dynamic environment permits. If the primary goal is educational, for instance, then the important points to be made must be incorporated into descriptions of the objects in which the browser is interested. The texts produced by the system are thus influenced by an *interaction* of the user's and the system's goals.

As the system is able to respond to the user's initiatives, it can be accurately termed **reactive**; however, it provides a rather different kind of reactivity than is normally discussed in the context of text planning systems. The most commonly discussed type of reactivity (see e.g. [Moore and Paris 1993]) is where a system allows the user the option to ask follow-up questions, to help clarify portions of the text which have not been fully understood. In such a case, the user has control over the way the system *executes* its goals, but not, ultimately, over the goals themselves. In the present system the situation is just the reverse: the user is able to determine the high-level goals about which objects are described, but the system is in control of the execution of these goals so as to include (if possible) information on its own agenda.

In the ILEX system, system goals are represented as a set of **messages** it desires to communicate. In the case of the Jewellery domain, these messages are seen as the educational goals of the display curator. Some messages from this domain are:

1. Jewellery need not be made of valuable materials.

2. Contemporary jewellers make innovative use of modern synthetic materials.

3. Jewellers frequently borrow techniques from other crafts such as pottery and weaving.

4. The distinction between 'art' and 'jewellery' is often blurred.

Attached to individual pieces of jewellery or classes of jewellery are canned stories that can be told if those items are chosen by the user. Each story in the story-database is tagged by the message or messages it works towards. In the future, stories will be ranked in relevance on the basis of the messages they serve, and will be expressed or not on the basis of this relevance.

If a number of stories have been given supporting a particular message, an explicit description of that message may be triggered; e.g., *Here again, we see that jewellery need not be made of valuable materials.*

The model of opportunism we are currently working with is very simple, but existing top-down NLG planning architectures do not seem to allow even this to be captured naturally. Support for opportunistic satisfaction of system goals is therefore an area in which further NLG research is required. We are currently looking to see what models of opportunistic planning (e.g. [Pryor and Collins 1994]) have to offer in this respect.

4 Conclusion

This paper has discussed several of the parameters of variation in dynamic hypertext systems. All of these parameters are associated with kinds of flexibility which seem to be desirable in the exploitation of the dynamic hypertext concept, but they are not necessarily independent or all useful together. Indeed, flexibility is not without cost; for instance, canned text can offer fluency which the most advanced NLG systems today have yet to reach.

Some of the parameters are relatively well-understood, and the state-of-the-art in NLG is sufficient to allow us to manipulate them effectively in dynamic hypertext systems. However, other parameters—particularly those involving user freedom and system goals—require further investigation.

The domain we have been investigating raises the particular challenge of mixed initiative dialogue: the visitor is in control of macro-level content selection, but the system selects lower-level content, with the twin goals of satisfying the visitor's curiosity, and of conveying key information. In a sense, meeting the challenge subverts the very idea of hypertext as we know it.

In conventional hypertext, the author selects content and connectivity, and then retires; the user is free to sample the hyper-document however they wish. Because of the author's disappearance, the effectiveness of the sampling is crucial: authors strive to place links and important content in all the right places, but some users still feel that they have to sample exhaustively every node of the hypertext. Nonetheless, hypertext is attractive because users value their freedom to sample as they choose.

In a system which opportunistically both satisfies user curiosity and its own informational goals, we can provide an ersatz author. However the user samples, they will always end up with the important content. In a sense, the users' freedom is an illusion: they cannot fail. When fishing on an ordinary river, an angler casts their line, and occasionally, if they do it well enough, and wait long enough, they catch a fish. We are proposing a new kind of river, in which an angler can cast their line haphazardly, and still pull out a perfect fish, time after time.

References

[Binsted et al. 1995] Kim Binsted, Alison Cawsey and Ray Jones. "Generating Personalised Patient Information Using the Medical Record", *Proceedings of AI in Medicine Europe*, 1995.

[Buchanan et al. 1995] Bruce Buchanan, Johanna Moore, Diana Forsythe, Giuseppe Carenini, Stellan Ohlsson and Gordon Banks, "An Intelligent Interactive System for Delivering Individualised Information to Patients", *Artificial Intelligence in Medicine* Vol 7, pp117-154, 1995.

[Carenini et al. 1993] Giuseppe Carenini, Fabio Pianesi, Marco Ponzi, and Oliviero Stock. "Natural language generation and hypertext access." *Applied Artificial Intelligence*, 7:135–164, 1993.

[Dale 1992] Robert Dale. *Generating Referring Expressions.* MIT Press, 1992.

[Dale and Milosavljevic 1996] Robert Dale and Maria Milosavljevic. "Authoring on Demand: Natural Language Generation in Hypertext Documents."In Proceedings of the First Australian Document Computing Conference, Melbourne, Australia, March 1996.

[Horacek 1995] Helmut Horacek. "More on Generating Referring Expressions." In *Proceedings of the Fifth European Workshop on Natural Language Generation*, pages 43–58, Leiden, The Netherlands, 1995.

[Kibby and Mayes 1988] M. R. Kibby and J. T. Mayes, "Towards Intelligent Hypertext". In McAleese, R., Ed., *Hypertext, Theory into Practice*, Intellect Books, 1988.

[Milosavljevic et al. 1996] Maria Milosavljevic, Adrian Tulloch and Robert Dale. "Text Generation in a Dynamic Hypertext Environment." In *Proceedings of the Nineteenth Australasian Computer Science Conference*, Melbourne, January 1996.

[Moore and Paris 1993] Johanna Moore and Cécile Paris. "Planning Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information." *Computational Linguistics* Vol 19, No 4, pp651-694, 1993.

[Paris 1993] Cecile L. Paris. *User Modelling in Text Generation.* Pinter Publishers, 1993.

[Pryor and Collins 1994] Louise Pryor and Gregg Collins, "Opportunities: A Unifying Framework for Planning and Execution", *Procs of the Second International Conference on Artificial Intelligence Planning Systems*, pp329-334, Chicago, IL, 1994.

[Reiter et al. 1992] Ehud Reiter, Chris Mellish and John Levine. "Automatic Generation of On-Line Documentation in the IDAS Project." In *Proceedings of the 3rd Conference on Applied Natural Language Processing*, Trento, Italy, April 1992.

[Reiter at al. 1995] Ehud Reiter, Chris Mellish and John Levine. "Automatic Generation of Technical Documentation". To appear in *Applied Artificial Intelligence*, 1995.

[Stotts and Furuta 1989] D. Stotts and R. Furuta, "Petri-net based Hypertext: Document Structure with Browsing Semantics", *ACM Transactions on Information Systems* Vol 7, pp3-29, 1989.