

An Architecture for Opportunistic Text Generation

Chris Mellish, Mick O'Donnell

Department of Artificial Intelligence, University of Edinburgh

80 South Bridge, Edinburgh EH1 1HN

Email: {chrism,micko}@dai.ed.ac.uk

Jon Oberlander, Alistair Knott

Human communication Research Centre, University of Edinburgh

2 Buccleuch Place, Edinburgh EH8 9LW

Email: {alisk,jon}@cogsci.ed.ac.uk

Abstract

In this paper we describe the architecture of the ILEX system, which supports *opportunistic* text generation. In web-based text generation, the system cannot plan the entire multi-page discourse because the user's browsing path is unpredictable. For this reason, the system must be ready opportunistically to take advantage of whatever path the user chooses. We describe both the nature of opportunism in ILEX's museum domain, and then show how ILEX has been designed to function in this environment. The architecture presented addresses opportunism in both content determination and sentence planning.

1 Exploiting opportunities in text generation

Many models of text generation make use of standard patterns (whether expressed as schemas (e.g. [McKeown 85]) or plan operators (e.g. [Moore and Paris 93])) to break down communicative goals in such a way as to produce extended texts. Such models are making two basic assumptions:

1. Text generation is goal directed, in the sense that spans and subspans of text are

designed to achieve unitary communicative goals [Grosz and Sidner 86].

2. Although the details of the structure of a text may have to be tuned to particulars of the communicative situation, generally the structure is determined by the goals and their decomposition. That is, a generator needs strategies for decomposing the achievement of complex goals into sequences of utterances, rather than ways of combining sequences of utterances into more complex structures. Generation is "top-down", rather than "bottom-up" [Marcu 97].

Our belief is that there is an important class of NLG problems for which these basic assumptions are not helpful. These problems all involve situations where semi-fixed explanation strategies are less useful than the ability to exploit *opportunities*.

WordNet gives the following definition of "opportunity":

Opportunity: "A possibility due to a favorable combination of circumstances"

Because opportunities involve *combinations* of circumstances, they are often unexpected and hard to predict. It may be too expensive or impossible to have complete knowledge about

them. Top-down generation strategies may not be able to exploit opportunities (except at the cost of looking for all opportunities at all points) because it is difficult to associate classes of opportunities with fixed stages in the explanation process.

We are investigating opportunistic text generation in the Intelligent Labelling Explorer (ILEX) project, which seeks automatically to generate a sequence of commentaries for items in an electronic catalogue (or museum gallery) in such a way as to reflect the interest of the user and also to further certain educational (or other) aims. The current domain of the system is the 20th Century Jewellery Exhibit in the Royal Museum of Scotland.¹ However, ILEX is designed to work with any domain where object descriptions are required. The key features of the ILEX application are:

- The system has an agenda of *communicative goals* to achieve, which reflect the goals of the curators.
- The user has the freedom to look at any object in the gallery at any time.
- The system produces a description of each object asked for by the user, such that:
 - each description contributes as best it can to the system’s goals;
 - the sequence of descriptions fits together into a coherent whole.

The result is a variety of *mixed-initiative dialogue*:

- the user is in control of the high-level communicative goal (what gets described);

¹The work is in collaboration with the National Museums of Scotland, Interactive Information and VIS Interactive Media.

- the system is in control of how the goal is realised (how the chosen object is described).

In such a dynamically unfolding environment, it is not possible to predict all possible paths through the interaction. The system must thus be ready to exploit opportunities in order to achieve its goals. In ILEX, the user’s arbitrary choice represents a horizon beyond which is it not practical to predict. Each generated page may be the last one to be generated and therefore has to be planned to achieve as much as possible on its own. Moreover, almost any part of the generated text can be optimised to exploit the arbitrary situation that the user has got themselves into.

2 Opportunities: evidence and models

2.1 Evidence: the goals of a museum curator

A museum curator seeks to achieve general educational goals through the description of a set of carefully selected objects. In general, the goals are to convey important generalisations (e.g. “Organic jewellery tends to have natural themes”) and to dispel important misconceptions (e.g. “Jewellery tends to be made of expensive materials”). These important points have to be brought in appropriately during the description of the exhibits which are selected by the visitor.

In order to see how a human being tackles such complex goals, we performed a “Curator of Oz” experiment, in which we chose an arbitrary sequence of exhibits in the 20th Century Jewellery gallery of the National Museum of Scotland and asked the curator to give us a commentary. The curator exploited opportunities of the following kinds.² Note that the classification here is only meant to

²In these examples, “V” indicates the visitor and “C” the curator.

be suggestive, and we don't claim that the categories are entirely disjoint or exhaustive.

Introducing general points/themes

suggested by the objects. The curator frequently moved from the objects to the general issues surrounding them, using the objects merely as an excuse to introduce these topics:

V: "There's a set of three objects here."

C: "What these symbolise for me are the preoccupations of the 1980's with..."

Moving to more interesting topics

Sometimes the curator indulged in much more extensive digressions:

V: "It's the suffragette brooch. Could you tell us something about that? ..."

C: "Suffragette jewellery is a subject which few people know about... The colours of the WSPU were purple, white and green... They had *sliced bread* wrapped in the colours..."

Reinforcing a point from the past

When an important point has already been made, an excuse to reinforce it by showing its relevance in a new situation is exploited:

V: "This one here..."

C: "Yes, you've made a link with the first piece that we looked at, which is the idea of a jewel which is also a work of art and a sculpture..."

Linking to previous items Making connections to items previously seen helps to improve the continuity of the discourse:

C: "...and it was work like this which directly inspired work like the Roger Morris brooch on the stand which we looked at earlier."

Suggesting future items Although the curator cannot control the choices of the visitor, nevertheless they can attempt to influence them in a way that will open up useful opportunities in the future:

V: "...object 9... Why is that there?"

C: "...That's there because I'm quite interested in refractory metals, which include titanium... there's another example in the same case ..."

is nothing like a conventional schema structure to the descriptions produced. The approach looks a lot more like putting together arbitrary pieces of interesting material subject to only very loose restrictions. This may not be the best way to produce a carefully-argued written text, and clearly the result is not always fluent according to stringent criteria. In some—but not all—respects, it resembles the unplanned discourses discussed by [Ochs 79]. Furthermore, in the interactive and relatively informal setting of a museum tour, it works.

We thus decided that ILEX should have a whole set of goals about things to say. These are linked into a single metalevel goal, which is something like "to achieve as many of the individual goals as possible, within the space available, in the context of a globally coherent discourse which maintains the reader's interest".

2.2 Models: planning for opportunities

We discussed above why top-down planning seems an unnatural basis for formulating an NLG model that can exploit opportunities.

There are, however, other models of planning that may be more appropriate.

ILEX is inspired loosely by ideas from *opportunistic planning* [Hayes-Roth and Hayes-Roth 79, Pryor 96]. Key elements of this are:

- Interleaving of planning and execution.
- Expanding “sketchy plans” as needed, taking into account the current state of the world.
- Flexible choice of tasks from an agenda.
- Recognition of opportunities through detection of *reference features*, followed by more thorough analysis.

Pryor’s work is implemented in PARETO, a planner for a simulated robot delivery vehicle. The vehicle is given orders to deliver various objects to various building sites, and needs to locate these objects at other sites. The system is opportunistic in that while the truck is working on one goal, it is always ready to switch to another if an object on its find-and-deliver list turns up. For example, if the truck stops at one place to pick up a hammer, it may notice a saw, which is also on its list, and thus pick it up and proceed to its delivery point.

Pryor’s planning occurs within a limited horizon—the robot only has certain knowledge in regards to the immediate location, and outside of that, the world is uncertain (objects are sometimes randomly moved between sites in the world).

ILEX inhabits a world analogous in certain respects to PARETO’s: each page is a site on the map, and it is up to us to find opportunities for realising our goals at each site. However, while in the truck world the system is in control of motion to the next site, in the museum, it is the user who chooses the next page. Conversely, while objects outside the truck’s immediate vicinity may move autonomously, for ILEX, facts and their values do not change.

Opportunistic planning has similarities with a number of other approaches to planning. It shares with *incremental planning* (used in NLG by [Cawsey 92]) the idea of starting to execute a plan before the plan is complete, and being prepared to repair the partial plan in the light of feedback. It shares with *reactive planning* the idea of being directed as much by the characteristics of the state of the world at execution time as by the pursuit of preconceived goals. However, unlike pure reactive planning it does acknowledge the need for explicit *plans* to be manipulated and it differs from many models of incremental planning in the extent to which the original plan can be diverted to exploit the characteristics of the world at execution time.

3 The ILEX architecture

To show how ILEX supports opportunistic text generation, we will here outline the parts of the system and the operation of its text planning. Basically the ILEX task agenda at each point consists of the facts that the system knows which have not yet been conveyed to the user. Each of these ‘tasks’ has an opportunity value (its educational value, assumed interest to the reader and contribution to coherence). At each point of the discourse, we ‘perform tasks’ (include facts) which provide the highest opportunity gain.

3.1 The Text Potential

The facts of our knowledge base are interconnected in various ways, and to facilitate content selection and structuring, we organise the facts into a *text potential* – a graph of facts interconnected in terms of thematic and rhetorical relations. The text potential is an intermediary stage between the knowledge base and text, motivated in a similar way to DRSs [Kamp 81] by the desire explicitly to represent the selection of possible knowledge structures that can be reflected linguistically. Most

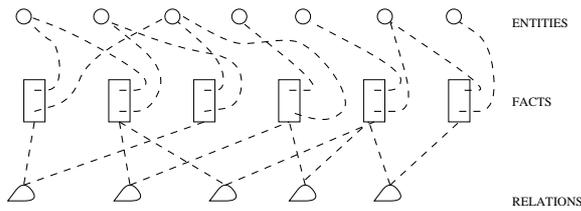


Figure 1: The Text Potential

of the text potential is precompiled, though some aspects of the text potential change dynamically and have to be computed on demand.

As Figure 1 shows, the text potential forms a three-tiered structure of *entities*, *facts* and *relations*. There are links between items in adjoining tiers, but no links within a tier or between entities and relations. We now discuss the three tiers in turn.

3.1.1 Entities

Entities are the participants in facts (things and qualities in terms of Penman’s Upper Model). Entities may be of two kinds: *specific* entities – such as an individual jewel or person; and *generic* entities – an entity representing some class of entities, such as Scottish jewellers, or art-deco brooches. Generic entities are treated essentially in the same way as specific entities in the text potential, for purposes such as the tracking of focus, anaphor generation, and so on.

3.1.2 Facts

Facts represent the relations between entities, in both events (e.g., X made Y), and states (e.g., X owns Y). In ILEX, we have assumed that all facts are binary (simple relations between two entities), e.g., `made-by(J-9999, King01)` represents the fact that the designer King made item J-9999. The binary assumption simplifies our architecture, allowing quicker text generation. At a later stage, we may allow more complex fact-representation.

Complex sentences can be formed through aggregating together these binary facts.

Each fact has the following fields:³

Pred : The name of the predicate connecting the two entities.

Arg1 : The entity in the relationship which the fact is primarily about. For instance, “J-999 was designed by Jessie King” is primarily about J-999, not about King.

Arg2 : The other entity in the relationship. This is sometimes another thing (such as “Jessie King”) and sometimes a quality.

Various other fields exist which detail the polarity, defeasibility, interest, importance and assimilation of the fact. Facts representing general principles or negations of general misconceptions are expressed using generic entities and can be included in a text just like any other facts.

3.1.3 Relations

Relation nodes represent relations between facts. Although based on conceptual relations, they qualify as rhetorical in that only the subset of relations that could explicitly be conveyed is included in the text potential. Relations include e.g., *Example*, *Concession*, *Amplification*, *Similarity*, *Contrast* and “*In that*”. Each relation has a nucleus and satellite (as in RST) as well as a set of precondition facts, which must be assimilated before the relation can be. There are no relations between relation-nodes in the text potential at present. Relation-nodes only link fact-nodes.

Relations in the text potential present a uniform interface as nodes connected to facts in the graph but are implemented by special-purpose procedures of varying complexity.

³Another type of fact node is used to contain canned text. The canned text is usually about associated with the focal object of the text, and no Arg2 field is provided.

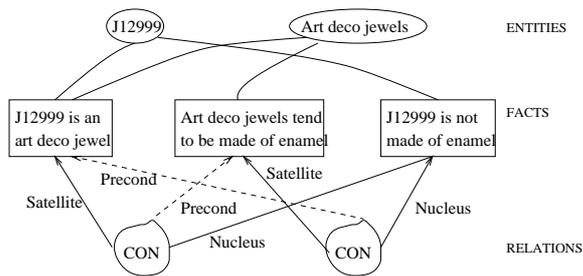


Figure 2: The Concession Relation

Figure 2 shows a small sub-graph of the text-potential, showing two *Concession* relations between facts.

3.2 Content Determination

ILEX plans a single page of text, describing a single entity, at a time. The text potential represents the information we can express, and the interconnectivity of information. When we receive the request for an entity description, the planner sets that entity as the global focus of the current page. Opportunistic planning then commences. The facts directly connected to that entity represent opportunities: the system can coherently include these facts in the text. If any of these facts are actually selected, then new opportunities are created in two ways:

- **Entity-based moves:** From the fact, we go to the argument which we didn't enter the fact from. We then select a new fact reachable from this node. See Figure 3. If we followed the Arg2 role of a fact, then we are in a sense selecting a new focus (local focus). The facts we generate about this entity should have the new entity as the focus. Thus in the example, King becomes the Theme of the second sentence. Sentences introduced using entity-based moves can be realised using an *Elaboration* relation to the starting fact.

An entity-based move from an individual

entity to its generic class entity can be made once the appropriate “isa” fact has been selected:

This item is an organic jewel.
Organic jewels tend to be ...

- **Relation-based moves:** from the initial fact, we follow a relation-node to some new fact. The new fact will be realised textually as a satellite to the original fact's nucleus. The type of the relation-node will determine the rhetorical relation of the link. See Figure 4.

Once we select a new fact in either of the ways described above, the new fact may act as the starting point for new opportunistic expansion. Alternatively, we may decide to backtrack to some earlier point, effecting a focus pop in Grosz and Sidner's [Grosz and Sidner 86] terms.

The selection of which opportunity to explore is determined by a number of heuristic factors. Firstly, facts are weighted according to the chain of relations back to the focus of the page [O'Donnell 97]. This is a way of preventing lengthly digressions from the supposed topic of the text. Secondly, each fact is associated with numbers which represent the opportunity 'value' of the fact. The opportunities are of two kinds:

Interest. the estimated value of the fact to the user, e.g. being made of plastic or paper are more interesting (*to the user*), because they are unusual in jewellery. Canned anecdotes about a piece of jewellery may also have high interest values.

Importance. the value of the fact as regards the system's educational agenda, e.g., the system considers it important to educate on stylistic development, so facts about styles are rated highly.

These values are moderated by a third fact annotation:

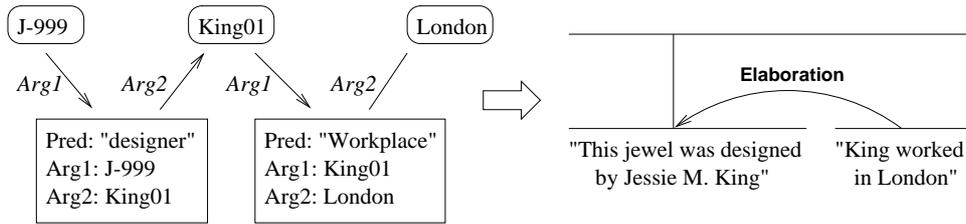


Figure 3: An Entity-Based Move

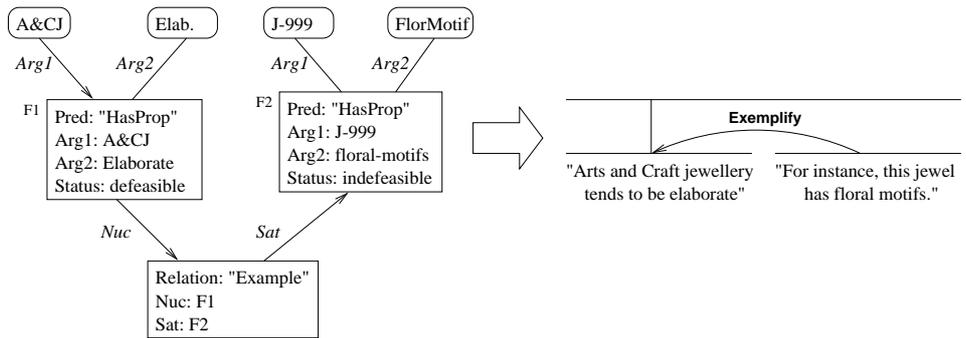


Figure 4: A Relation-Based Move

Assimilation. the degree to which the fact is assumed known to the user, either from general knowledge, or through prior mentions in the web interaction (these values change dynamically).

These three values, interest, importance and assimilation are taken together to calculate the opportunity value of each fact, which is used together with the evaluation of the chain of relations to select which textual opportunities to follow. We have no special theory about where interest and importance come from, though the above examples suggest that there may be domain- and user-type-specific rules that can be used to derive some of them.

In summary, content-determination in ILEX is seen as the task of optimising the selection of opportunities that are offered by the topic of the text, subject to not moving too far from that topic. The result of content-determination is a connected subgraph of the text potential (Figure

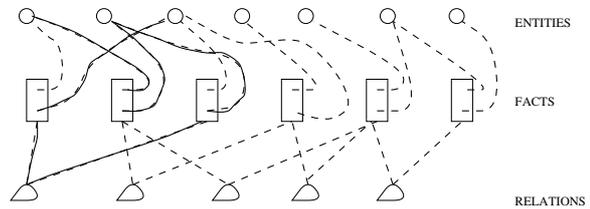


Figure 5: Result of Content Determination

5). The use of interest and importance in ILEX is analogous to the use of “salience” in [McDonald and Conklin 82]. Because the process is seen as a graph traversal problem, there are also similarities with work on generating text from semantic networks [Simmons and Slocum 72, Sibun 92]. In a sense, our work aims to combine the best of both.

3.3 Text Planning

Although the process of content determination has worked through a number of moves

that may be made in the generated text, the result is not the kind of tree structure that one needs for realisation and also has been influenced only by local considerations of coherence. Text planning therefore requires the following two steps:

1. Extend the subgraph to a complete subgraph that includes *all* the relations linking the selected fact nodes.
2. Produce from this an “optimal” selection of relations, so as to give rise to an RST tree including all the selected facts.

Once the complete subgraph has been obtained, the text planning problem is exactly that described by [Marcu 97]. The idea of combining a set of facts together into an “optimal” text is also compatible with [Hovy 90] and the earlier work of [Mann and Moore 81]. Again this involves exploiting opportunities. For instance, in order to avoid an awkward focus shift at some point, one might attempt to include a selected fact about a new entity immediately after another one that mentions the same entity. Other text planning operations that are opportunistic in nature include aggregation [Dalianis and Hovy 96] and redundancy suppression [McDonald 92], though we will not consider these here.

ILEX has been designed in a modular way such that each module (processing stage) can be easily pulled out and a replacement module inserted. This is true for text planning, where we are currently experimenting with a range of planning algorithms. Again, these are all opportunistic in nature, rather than being strongly goal-directed or schema-based. We could use Marcu’s methods directly, but are exploring more widely because:

- We would like to take into account a wider range of preference criteria, some of which involve global properties of trees (e.g. preferences based on focus and on sizes of substructures). (Marcu in fact uses a global evaluation based on “right

skew” in his work on rhetorical *parsing*.) We would like to develop global criteria further.

- We argue elsewhere that entity-based elaborations are rather different from other rhetorical relations and that the algorithms and representations should reflect this directly. (See [Knott et al 1998].)
- Marcu’s approach involves finding an optimal solution to a constraint satisfaction problem and enumerating all RST trees compatible with a given sequence of facts. We believe that the combinatorics of this will be unattractive for large examples (since constraint satisfaction is intractable in the general case) and wish to investigate heuristic approaches.

One of our current text planning algorithms uses a deterministic procedure to map a sequence of facts onto a single RST tree and a genetic algorithm to search for a sequence whose tree is as highly-valued as possible [Mellish et al 1998]. This is not yet integrated with the main ILEX system, but when run on content selected by ILEX produces text plans which could be realised as texts such as that shown in Figure 6. Note that for this text the realisation (including aggregation and referring expression generation) has been done by hand, though the ordering and choice of rhetorical relations is performed by the system. Although there are perhaps some places where limited schemas would have helped this text, nevertheless the system has been quite successful in interleaving more “important” facts about the designer and the style with facts about the topic of the text.

3.4 ILEX and Opportunistic Planning

With this description of ILEX in mind, we can explore the analogy with PARETO in more detail. Where PARETO embarks on the

This item is made from enamels. It was made by King. King was a Scottish designer who lived in London. This jewel is an Arts and Crafts style jewel. It is a pendant-necklace and was made in 1905. This jewel is made from yellow metal and moonstones. It has festoons. This jewel has an elaborate design: it has floral motifs. This jewel was produced in limited quantity and is an example of craft jewellery. This jewel is set with jewels: it features cabuchon stones. It is worn round the neck. This jewel is very flexible. Indeed, Arts and Crafts style jewels are usually flexible. Arts and Crafts style jewels usually have an elaborate design, feature cabuchon stones, are produced by a single craftsman and use oval stones.

Figure 6: Example output text (realised by hand)

execution of a sketchy plan to start moving around the truck world, ILEX embarks on a graph traversal, starting out from the topic entity and guided by the desire not to digress excessively. The text potential offers options to ILEX in a similar way to PARETO's world. In PARETO, reference features indicate possible opportunities; in ILEX this role is played by the interest and importance annotations. Deeper analysis is required by PARETO before seizing an opportunity; this is probably analogous to the way that ILEX attempts to find the globally best way of incorporating material into the RST tree.

4 Conclusions

An *opportunistic* planning algorithm seems to be required for the ILEX domain (and some other domains too). ILEX certainly does have goals – to educate the user and keep them interested. But these are essentially compiled away into the content of the text potential (e.g. the entities and general facts that are

included) and the interest and importance annotations on facts. At runtime, ILEX just has the task of selecting the best opportunities, given the situation that the user has reached. These opportunities arise not just in how content is selected but also in how it is structured. We conclude that there a strong analogy between the way that ILEX operates and the techniques used in opportunistic approaches to planning in other domains, and this connection is worthy of further exploration.

5 Acknowledgements

The ILEX project is supported by EPSRC grant GR/K53321, and the third author is supported by an EPSRC Advanced Fellowship. We acknowledge the valuable assistance of the National Museums of Scotland.

References

- [Cawsey 92] Cawsey, A., *Explanation and Interaction*, MIT Press, 1992.
- [Dalianis and Hovy 96] Dalianis, H. and Hovy, E., “Aggregation in Natural Language Generation”, in Adorni, G. and Zock, M., Eds., *Trends in Natural Language Generation*, pp88–105, Springer Verlag, 1996.
- [Grosz and Sidner 86] Grosz, B. and Sidner, C., “Attention, intentions and the structure of discourse”, *Computational Linguistics*, **12**, 175-204, 1986.
- [Hayes-Roth and Hayes-Roth 79] Hayes-Roth, B. and Hayes-Roth, F., “A cognitive model of planning”, *Cognitive Science*, **3**, 275-310, 1979.
- [Hovy 90] Hovy, E., “Unresolved issues in paragraph planning”, in Dale, R., Melliish, C. and Zock, M., (eds.) *Current Research in Natural Language Generation*. Academic Press, 1990.

- [Kamp 81] Kamp, H., “A theory of truth and semantic representation”, in Groenendijk, J. A. G., Janssen, T. M. V. and Stokhof, M. B. J. (eds.) *Formal Methods in the Study of Language* **136**, pp.277–322. Amsterdam: Mathematical Centre. Tracts.
- [Knott et al 1998] Knott, A., Oberlander, J., Mellish, C. and O’Donnell, M., “Non-local dependencies in discourse structure tree: the case of the ELABORATION relation”. Ms.
- [Marcu 97] Marcu, D., “From local to global coherence: a bottom-up approach to text planning”, in *Proceedings of AAAI-97*, American Association for Artificial Intelligence, 1997.
- [Mann and Moore 81] Mann, W. and Moore, J., “Computer generation of multiparagraph english text”, *Computational Linguistics*, **7**, 17–29, 1981.
- [McDonald 92] McDonald, D., “Type-Driven Suppression of Redundancy in the Generation of Inference Rich Reports”, in Dale, R., Hovy, E., Rösner, D. and Stock, O., Eds., *Aspects of Automated Natural Language Generation*, pp73–88, Springer Verlag, 1992.
- [McDonald and Conklin 82] McDonald, D. and Conklin, E. J., “Salience as a simplifying metaphor for natural language generation”, in *Proceedings of AAAI-82*, American Association for Artificial Intelligence, 1982.
- [McKeown 85] McKeown, K., *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*, Cambridge University Press, 1985.
- [Mellish et al 1998] Mellish, C., Knott, A., Oberlander, J., and O’Donnell, M., “Experiments using stochastic search for text planning. Ms.
- [Moore and Paris 93] Moore, J. and Paris, C., “Planning texts for advisory dialogues: capturing intentional and rhetorical information”, *Computational Linguistics*, **19**, 1993.
- [Ochs 79] Ochs, E., “Planned and unplanned discourse”, in Givón, T. (ed.) *Syntax and Semantics 12: Discourse and Syntax*, pp.51–80. New York: Academic Press.
- [O’Donnell 97] O’Donnell, M., “Variable-length on-line document generation”, in *Proceedings of the 6th European Workshop on Natural Language Generation* Gerhard-Mercator University, Duisburg, Germany, 1997.
- [Pryor 96] Pryor, L., “Opportunity recognition in complex environments”, in *Proceedings of AAAI-96*, American Association for Artificial Intelligence, 1996.
- [Sibun 92] Sibun, P., “Generating texts without trees”, *Computational Intelligence*, **8**, 102–122, 1992.
- [Simmons and Slocum 72] Simmons, R. and Slocum, J., “Generating English discourse from semantic networks”, *Communications of the ACM* **15**, 891–905, 1972.