

Sentence Ordering in Electronic Navigational Chart Companion Text Generation

Julie Sauvage-Vincent

Institut Mines-Télécom
Télécom Bretagne
Technopôle Brest Iroise
CS 83818
29238 Brest Cedex 3, France
jsauva01@telecom-bretagne.eu

Yannis Haralambous

Institut Mines-Télécom
Télécom Bretagne
Technopôle Brest Iroise
CS 83818
29238 Brest Cedex 3, France
yannis.haralambous@telecom-bretagne.eu

John Puentes

Institut Mines-Télécom
Télécom Bretagne
Technopôle Brest Iroise
CS 83818
29238 Brest Cedex 3, France
john.puentes@telecom-bretagne.eu

Abstract

We present the sentence ordering part of a natural language generation module, used in the framework of a knowledge base of electronic navigation charts and sailing directions. The particularity of the knowledge base is that it is based on a controlled hybrid language, that is the combination of a controlled natural language and a controlled visual language. The sentence ordering process is able to take into account hybrid (textual and visual) information, involving cartographic data, as well as landscape “read” by the navigator.

1 Introduction

The French Marine Hydrographic and Oceanographic Service (SHOM, *Service Hydrographique et Océanographique de la Marine*) issues, on a quadrennial basis, *Instructions nautiques*, a series of nautical books providing navigators of coastal and intracoastal waters with useful information.

Instructions nautiques are intended as a complement to Electronic Navigational Charts (ENCs) and add a wide variety of essential information not provided in the ENCs for maritime navigation. In this sense they are considered as *companion texts* of ENCs.

Information found in *Instructions nautiques* are in some cases subject to real-time updates. To make this possible, an ongoing SHOM project is to build a knowledge base (KB) covering both ENCs and nautical instructions. This KB is intended to communicate with ENCs and more globally with any compatible Electronic Charts Display Information System.

Updates are planned to be operated mainly by SHOM domain experts, who may not be necessarily proficient in ontology formalism or in language technology. Therefore, it has been decided

to use a *controlled natural language* for exchanges between experts and the KB (Haralambous et al., 2014). On the other hand, information contained in the KB covers not only (textual) *Instructions nautiques* but also (visual) ENCs. These two modalities are tightly bound, coreferential and complementary: each modality covers information that the other is unable to transmit.

In order to establish intermodal coreferentiality and complementarity, a new type of controlled language has been defined (Haralambous et al., 2015), called *controlled hybrid language* (CHL), which is intended to be based on hybrid sentences, like for instance:



[The wreck of Morania 130]
lies at the bottom of [lake
Erie].

In Fig. 1 (on the next page), the reader can see this (multimodal) sentence analyzed. On the bottom of the figure one can see the two visual and textual modalities; and above them, the corresponding syntactic trees: on the right, the usual constituency syntax tree of the textual sentence (georeferenced named entities, placed in brackets, are considered as indivisible noun phrases); on the left, the syntax tree of a small part of the map, considered as a sentence in a visual language, using the Symbol-Relation formalism (Ferrucci et al., 1996; Ferrucci et al., 1998). In both cases, the formal grammars have synthesized attributes (in the sense of Knuth (1968)) carrying semantics: using a bottom-up synthesis approach we obtain their semantics, represented as First-Order Logic graphs (predicates are hexagons, connectors are circles, functions are rounded rectangles, and constants are rectangles). Once the two graphs are established, and after a coreference resolution step, they are transformed and merged into the KB graph, at the top.

When starting from the KB, operators \mathfrak{V} and \mathfrak{T} filter their input into information that is represented visually and information that is represented

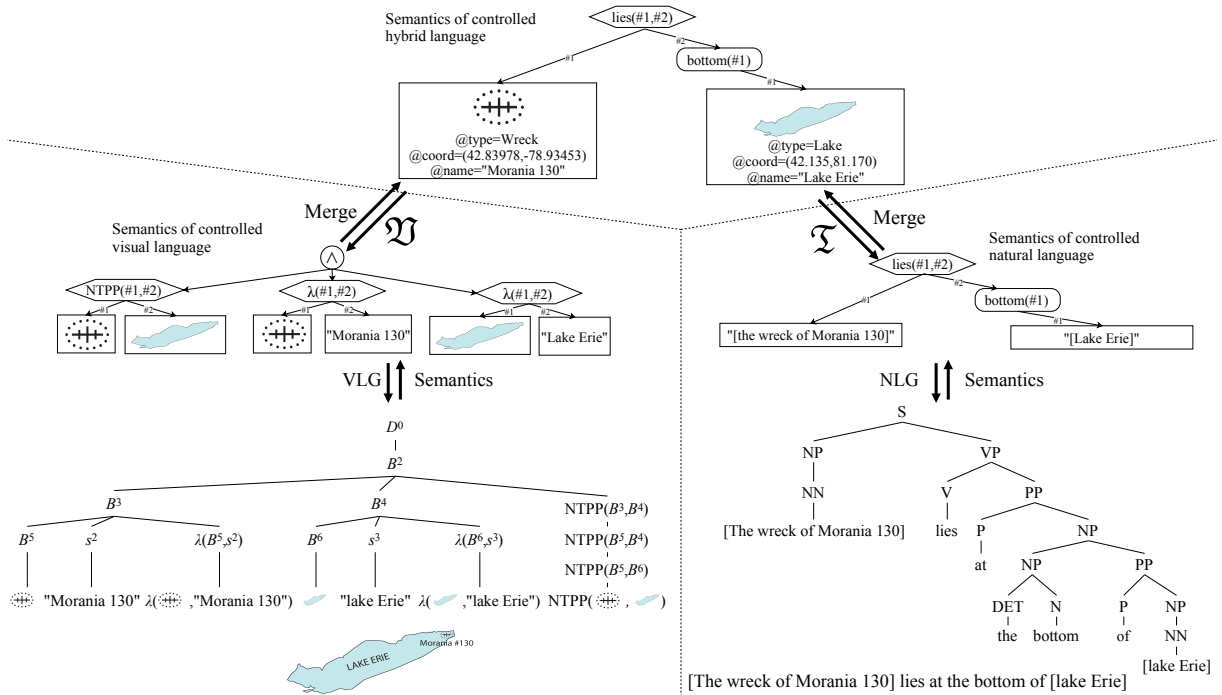


Figure 1: An example of sentence in the CHL INAUT (taken from Haralambous et al. (2015)).

textually (with some redundancy in order to establish coreferential entities). Their outputs are graphs corresponding to FOL formulas. To obtain text, we use NLG and VLG (visual language generation) to obtain a part of the map.

The goal of the INAUT NLG module—which we are currently developing—is to produce the most *fluent*¹ multi-sentence texts possible. This paper addresses the stage of *sentence ordering* (as part of discourse planning), which plays a central role in the achievement of this goal.

2 Related work

Ordering sentences to create a natural and understandable paragraph for the reader is part of what Reiter and Dale (2000) call *discourse planning*.

A widely used approach to discourse planning is based on rhetorical structure theory (Mann and Thompson, 1988), which requires writing a rule for each textual structure. Although this solution has been proved efficient in various contexts (cf. Taboada and Mann (2006)), this is not the case for the *Instructions nautiques* corpus, written by different authors who do not necessarily share the same rhetorical structures and processes.

The NaturalOWL (Androutsopoulos et al. (2013)) system per se could be used for INAUT

¹In the sense of criteria S1–S5 of Androutsopoulos et al. (2013, p. 703).

automatic text generation, if there were not for some major differences. NaturalOWL is essentially based on Centering Theory, i.e., it respects thematic intersentential coherence. In our case there are some additional issues, related to the fact that INAUT is build upon a *hybrid* language: information contained in text is not the only input anymore, and we must guarantee conformance to the itinerary of a vessel, to the geographic “guiding path” of each *Instructions nautiques* volume and, last but not least, to the visual characteristics of the landscape. Indeed, *Instructions nautiques* are, inter alia, textual interpretations of the real world as seen by the navigator, and for this reason sentence order must respect the order navigators “read” the landscape. Another major difference in our system is real-time interaction with users. The latter necessarily has an impact on the structure of generated text: when content determination may be relaunched on different data every few milliseconds, the stability of generated text becomes a major issue.

3 Data and pre-processing

The corpus consists of 462 INAUT controlled hybrid language sentences manually translated from the legacy *Instructions nautiques*.

Let us consider the first step of NLG, namely content determination.

3.1 Content determination

Among the attributes of nodes in the KB we have coordinates for all geolocalized objects. Therefore, hybrid language structure provides a link between geolocalization and (textual) sentence entities. Content determination can be initiated by both (1) textual criteria (selecting a paragraph in the document tree structure), and (2) visual criteria (selecting an area on an ENC).

In case (1), one obtains immediately a subgraph of the KB by taking the nodes hierarchically located under the chosen paragraph node. In case (2), a query sent to the KB server returns all georeferenced nodes located entirely or partially in the selected area of the map. In both cases one obtains a (not necessarily connected) subgraph of the KB.

By the nature of the data, two further steps are needed, both obtained by inference, but on different kinds of data, namely spatial data and temporal/meteorological context.

The first inference step concerns cases where information about a geolocalized object can be inferred from the map. More generally, one can extract knowledge from the map data, which will complement, enhance, or contradict the textual data.

As for temporal and meteorological context, tide and weather conditions obviously have an impact on navigation. This is also the case for regulations based on a schedule. Inference based on these data may act as a filter on the subgraph obtained either hierarchically or by area selection.

Finally, an important feature of the INAUT system is to inform navigators on potentially dangerous situations. By attaching—either manually or by applying inference to geography and context—a *dangerousness coefficient* to specific nodes under given conditions, the system may introduce specific warnings in the generated text.

4 Modelling the domain experts sentence ordering process

We consider the discourse planner as a multicriteria decision process based on frequent patterns of the writing process. Therefore, our main task is to model the implicit knowledge of authors concerning the description of a maritime environment.

4.1 Domain experts sentence ordering process

We detected common patterns in the way authors describe the maritime environment, and will try to discuss them from a cognitive and linguistic point of view.

These patterns are constrained by several criteria: our approach is to assign score to each criterion found in a sentence, in order to calculate the global sentence score in our “bag” of sentences, and reorganize the latter by sorting it in decreasing order of score. The greater the score, the greater the likelihood for the sentence to appear at the beginning of a paragraph. The computation of the score is done by the sum $f(s) = \sum_{i=1}^n c_i \cdot w_i$ where s is a sentence, c_i is a criterion value and w_i is the corresponding score. Given a set S of n sentences s_i , if $f(s_1) > f(s_2)$, then the sentence s_1 is more likely to precede s_2 .

To assign score to objects, we must understand which features domain experts use to describe a natural environment in general.

Let us consider the different features used in our ordering sentences module.

Landmarks When dealing with (a) authors tend to use *landmarks* as much as possible. Selection of elements useful in assisting human navigation in an open space has been addressed in the context of urban orientation. Michon and Denis (2001) attest the landmark usage preference in order to identify areas where difficulties in term of way finding are likely to occur.

We find this preference in our corpus as well: *Instructions nautiques* authors often prefer man-made landmarks—that facilitate the environment reading—over natural objects.

Geometric primitives Objects occurring in the description of a map or of a landscape can be of three different topological natures: *areas*, *lines* and *points*. We observed that SHOM domain experts describe objects in this order: *polygonal shapes* before *lines*, before *points*. According to Brosset et al. (2008) this can be explained by the fact that, from the point of view of observers, natural environment is seen as a *spatial network*: linear objects structure the network with edges and links, polygonal shapes act as a partition of the space, and finally points act as visual landmarks.

Name and size Two other features are directly connected to individual objects: their *size* and *name*. Indeed, named objects appear more frequently in the corpus than unnamed ones and larger objects more frequently than smaller ones.

Proximity spaces Another feature taking part in the multicriteria decision process is *geographic position relative to the vessel*.

When receiving directions, users tend to create by anticipation a mental representation of the route—whether they are standard or problematic routes. Unlike pedestrian navigation, maritime navigation requires a most precise representation of the surrounding and forthcoming environment.

According to Tversky (2003) humans structure environment in various mental spaces. Le Yaouanc et al. (2010) extended Tversky’s spaces to *proximity spaces*. These structure the visual perception of the landscape and therefore, logically, also its description. Proximity spaces are defined by actions users are able to perform within them. We distinguish four different proximity spaces (from the closest to the observer to the furthest away): (a) the *space of the body*, (b) the *experienced space*, (c) the *distant space* (d) and, finally, the *space at the horizon*. In their paper, Le Yaouanc et al. (2010) state that the different subjects of their study have used an order following these proximity spaces when describing an environmental scene.

It is interesting to note that in the SHOM corpus the order assigned by domain experts is the reverse of the one stated above in 93% of the cases. This difference relates to the fact that Le Yaouanc et al. (2010) used terrestrial environmental scene descriptions while the SHOM corpus deals exclusively with offshore environmental scene descriptions.

Thus, the further away objects are, the greater the score assigned to them. *Proximity spaces* are a typical example of an *hybrid feature*: the textual part alone would be clearly insufficient in providing information about size and position of objects.

Cardinal directions In the same spirit, we add yet another feature, namely *cardinal directions*. Indeed the latter provide an additional hint on the order of sentences in a paragraph since an environmental scene is usually observed in the reading direction of the observer (Nachson and Hatta, 2001; Fuhrman and Boroditsky, 2010), in our case from

left to right for 84.8% of the paragraphs where the description of objects is done in a longitudinal way.

Using the various features mentioned in this section, we have built a SVM classifier for ranking sentences. The classifier provides a lattice structure of ranked sentence pairs. Out of this lattice we obtain a best possible global order of sentences by a standard lattice-traversal algorithm.

4.2 The Stability Issue

Content determination, as part of the NLG process, depends on several parameters (the area selection, the temporal and meteorological context, etc.) which operate on three different temporal scales affecting NLG: slow landscape changes imply very few KB updates but temporal and meteorological context changes may need to be updated several times daily. Finally, selection updates done on the GUI with a mouse may be only milliseconds apart.

All three temporal scales, and the last one at the highest degree, raise the problem of *NLG stability*: a text should not change while the user is reading it or while the reader is using the mouse to change the selection area.

The issue of stability is a general NLG issue, and as such also affects sentence ordering. Changing the sentence order of a paragraph can be extremely disturbing for the reader.

In fact, user interaction with the GUI causes not only visual changes, but also simultaneous multi-level linguistic structure changes. To overcome this issue we introduce the method of *smooth text generation*, as follows:

We consider the function T that maps the values of the various text generation parameters to the text generated. This discrete function is “smoothed” in the following way:

1. When the mouse crosses a boundary between two areas covering the same nodes but different sentence orders, then the same sentence order is kept, until some nodes disappear or new nodes appear.
2. When the mouse enters a zone covering new nodes, then the sentences generated out of these nodes are—as much as possible—added at the end of the generated paragraph.
3. Generated text updates are slightly delayed so that a quick mouse move will not alter the gen-

erated text until the mouse is still for a time duration longer than a given threshold.

5 Conclusion and Future Work

We presented in this paper the sentence ordering part of the natural language generation module of the INAUT system. The particularity of this system is that it is based on a controlled hybrid language and hence covers simultaneously textual and visual knowledge.

We have shown that hybrid features (textual and visual) can be used to build a classifier that orders sentences in a paragraph.

Future work in the project involves a two-parts evaluation —(1) an automatic method based on comparison with the legacy corpus, and (2) a human-centered evaluation— and the exploration of other hybrid features impacting on sentence order, in particular by using the domain experts feedback of the second evaluation phase.

Furthermore, we will also consider hybrid language generation, i.e., having the system choose which information will be represented in visual or in textual modality, and insure coreferential redundancy among the modalities.

References

- Ion Androutsopoulos, Gerasimos Lampouras, and Dimitrios Galanis. 2013. Generating natural language descriptions from OWL ontologies: the NaturalOWL system. *J. Artif. Int. Res.*, 48:671–715.
- David Brosset, Christophe Claramunt, and Éric Saux. 2008. Wayfinding in natural and urban environments: a comparative study. *Cartographica*, 43(1):21–30.
- Filomena Ferrucci, Giuliano Pacini, Giorgio Satta, Maria I. Sessa, Genoveffa Tortora, Maurizio Tucci, and Giuliana Vitiello. 1996. Symbol-Relation Grammars: A Formalism for Graphical Languages. *Information and Computation*, 131:1–46.
- Filomena Ferrucci, Genny Tortora, Maurizio Tucci, and Giuliana Vitiello. 1998. Relation Grammars: A Formalism for Syntactic and Semantic Analysis of Visual Languages. In Kim Marriott and Bernd Meyer, editors, *Visual Language Theory*, pages 219–243. Springer.
- Orly Fuhrman and Lera Boroditsky. 2010. Cross-cultural differences in mental representations of time: Evidence from an implicit nonlinguistic task. *Cognitive Science*, 34(8):1430–1451.
- Yannis Haralambous, Julie Sauvage-Vincent, and John Puentes. 2014. INAUT, a Controlled Language for the French Coast Pilot Books *Instructions nautiques*. In Brian Davis, Kaarel Kaljurand, and Tobias Kuhn, editors, *Controlled Natural Language*, volume 8625 of *Lecture Notes in Computer Science*, pages 102–111. Springer.
- Yannis Haralambous, Julie Sauvage-Vincent, and John Puentes. 2015. A hybrid (visual/natural) controlled language. Submitted.
- Donald E. Knuth. 1968. Semantics of context-free languages. *Math. Syst. Theor.*, 2:127–145.
- Jean-Marie Le Yaouanc, Éric Saux, and Christophe Claramunt. 2010. A semantic and language-based representation of an environmental scene. *Geoinformatica*, 14(3):333–352.
- William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.
- Pierre-Emmanuel Michon and Michel Denis. 2001. When and why are visual landmarks used in giving directions? In Daniel R. Montello, editor, *Spatial Information Theory*, volume 2205 of *Lecture Notes in Computer Science*, pages 292–305. Springer.
- Israel Nachson and Takeshi Hatta. 2001. Directional tendencies of Hebrew, Japanese, and English readers. *Perceptual and Motor Skills*, 93:178–180.
- Ehud Reiter and Robert Dale. 2000. *Building Natural Language Generation Systems*. Studies in Natural Language Processing. Cambridge University Press, Cambridge, U.K.
- Maite Taboada and William C. Mann. 2006. Applications of rhetorical structure theory. *Discourse studies*, 8(4):567–588.
- Barbara Tversky. 2003. Structures of mental spaces. how people think about space. *Environment and behavior*, 35(1):66–80.