Creating knowledge maps using Memory Islands

Bin Yang, and Jean-Gabriel Ganascia

LIP6 laboratory, University Pierre and Marie Curie, Paris, France
{bin.yang, jean-gabriel.ganascia}@lip6.fr

Abstract. Knowledge maps are useful tools for managing and sharing the large-scale hierarchical knowledge that have recently started to be widely applied. In this paper, we discuss how knowledge maps can be generated by using Memory Islands. Memory Islands is our work on progress visualization technique, which was inspired by the ancient "Art of Memory". It consists of automatically creating the spatial cartographic representation of a given hierarchical knowledge (e.g., ontology) and with the help of its interactive functions, the users can navigate through the artificial landscape, to learn from the knowledge and retrieve them. We present some preliminary results of representing different hierarchical knowledge to show how the knowledge maps created by our technique work. We would like to discuss with the Information Retrieval community about interactive functions such as searching.

Keywords: Information Visualization, Knowledge Visualization, Ontology, Information Discovery

1 Introduction

Recently, many researchers begin interdisciplinary cooperation in emerging scientific research areas such as Digital Humanities and Digital Libraries, who rely on visualization techniques to help them develop their domain with their information tasks by use of both abstract data structures and information contents (e.g. ontology or encyclopedia). Knowledge maps are one of the promising tools for visualizing knowledge, as it could help the users to access the knowledge contents with an overview of the knowledge structures. With a knowledge map, the users can navigate and have an insight (given by the designer of the map or from the visualization technique) on the knowledge, as well as perform search tasks and share tasks over the same domain.

In this paper, we briefly express how the knowledge can be mapped with the Memory Islands. Memory Islands is a cartographic visualization technique, which was designed to amplify the advantages brought by geographic metaphors. It provides with interactive functions for navigation inside the knowledge map. It also tries to bring some advantages to the users for helping them with their information tasks with both abstract data structures and information contents. In the following we will present some of the knowledge maps we have generated with Memory Islands, discuss our first attempts with the knowledge orders and the advantages of creating knowledge maps with Memory Islands.

2 Map Based Visualization and Memory Islands

Map metaphors are well used in information visualization to create knowledge maps to help the user of the knowledge, Thorndyke et al. have proved the advantages of learning from a map [7]: the knowledge could reside in the memory like a physical map, which improve the user's capacity for memorizing the knowledge. Gansner et al. [4] summarized a number of existing map-based approaches. Skupin et al. [6] have summarized some existing spatialization methods (e.g., by applying tessellation).

Memory Islands is a visualization technique and our current work in progress; it was inspired by the ancient "Art of Memory" technique, which describes how people in the antiquity and the Middle Age used spatialization to increase their memory capacity. More specifically, our spatialization algorithms are inspired by the method of "loci" (plural of "locus", Latin for place or location) in the "Art of Memory"; it consists of creating a virtual map and associating each entity to designated areas on the map. Geographic metaphors such as distance, colors, label's size, and type of points, have been introduced to improve the expressivity of the underlying knowledge. We transform structured knowledge (e.g., trees or ontologies) into a cartographic representation in which each concept is associated to specific points on the map (Fig. 1). We are employing labeling algorithms to decide the zoom level we need in order to display all the knowledge contents in the result representations without their overlapping.

Therefore with our technique, even large knowledge could be represented in an interactive knowledge map, which allows the user to navigate through the virtual map (island), with borders, paths and boundaries. According to the research of Darken et al. [3], with these features, the users greatly benefit navigation performance.

The generation of the proxy elements (landscape) can be done either by applying clustering techniques directly to the original data space or based on created spatial 2D distributions. However using straight clustering techniques makes it harder for the users non-experienced to navigate in ontologies, making their visualization not as useful as one might think. Our spatialization algorithms generate a virtual geographic map (island or polygon) from the data enabling the visualization of both its structure and information content; the users could therefore benefit from the geographic metaphors which were used to design our visualization tool to achieve their information tasks.

2.1 Geographic Metaphors

Beside the landscape metaphor, we introduce geographic metaphors to express the hidden meaning inside the data. Here is a brief description of the metaphors we used in our Memory Islands.

- Proportion: like shown in Fig.4, we use the proportion metaphors to map the number of articles in the encyclopedia to the size of its area in the island.
- Distance/centrality: the centrality point will be the most important concept in the ontology and will serve as basis to calculate the distance from one concept to its subconcepts. The distance (ordering) between the nodes in the same hierarchical level with the same parent could be measured by the semantic similarity measures.

- Label's Font sizes / Point type: the label's font sizes will show the importance of a concept (e.g., level in a taxonomy), while the point type will describe a node's type in ontology in result island.
- Colors: the concept's area is colored according to its level and the type of its relation with its parents in the structure.

2.2 2D or 3D Representation

To design a cartographic representation, there are always choices between the representations in 2D and 3D. We have the following two main reasons to stay with 2D representation: firstly, the users are familiar with the 2D map services (like the Google Map), and learning from 2D map has already proved its advantages for memorization. According to Cockburn [2], he strongly suggests that "the effectiveness of spatial memory is unaffected by the presence or absence of three-dimensional perspective effects in monocular static displays." Therefore he states adding a third dimension would only increase the navigation costs.

2.3 Interactive Functions

With the help of the interactive function (zoom) and the label placement algorithms [8], we could visualize both the information contents and the structure. Here, we will discuss with the interactive functions of our technique including the current supported function and our planed interactive function:

- Pan and Overview: we provide an overview map to show the current position and overview structure of the knowledge, while with the pan function, users can easily navigate in virtual reality.
- Zoom: it helps focus on one part of information on the island.
- Details on demand: clicking a label or a point, a panel will appear and the user can access the sources (e.g., encyclopedia) of the concept. Also paths can emphasize the connections between related points based on the relations between the concepts.
- Search and Filter: visualization needs to provide a small search engine; with it the users can search for the information associated with the concept, such as instances. An ontology based filter function will also be added at a later stage.
- Sharing and collaborative update: we are currently working on the design of this function to extend our technique to support collaborations. The user could share his own experiences (e.g., exploration history) with Memory Islands or engage in collaborative development of knowledge for Synergistic Learning.

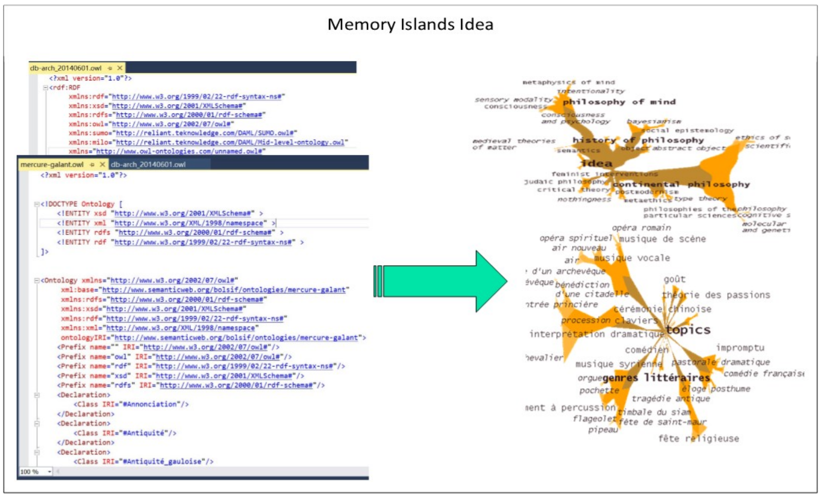

Fig. 1. Memory Islands Technique: transform hierarchical structure knowledge (ontology) into a 2D space representation.

3 Different Knowledge Maps created with Memory Islands

We will show in this section some knowledge maps that were created with our visualization technique, and how the users can navigate through by the interactive functions provided. If a knowledge could be presented in a formal ontology or a tree-structure could be found from that underlying knowledge, it could be visualized by our Memory Islands technique.

3.1 Text and Document

We use an original unsupervised approach for Named Entity Recognition and Disambiguation (UNERD) [5] with a French knowledge-base and a statistical contextual disambiguation technique that slightly outperformed Stanford's NER Classifier (when trained on a small portion of manually annotated data). It helps us visualize some children books for such entities as People, Locations or Organizations. We try to build a knowledge structure based on the table of contents with the named-entities recognized in each chapter (Fig. 2).

A second example is a representation of a public debate (Fig. 3) in the French Parliament - the National Assembly. Using data from the documents we can link debate topics with different speakers as well as cluster sub-topics around the main topics.

Currently we are experimenting with the representation of extracurricular books and documents to help the students enhance their learning. The objective of our method is to visualize the contents, according to the structure of the represented documents.

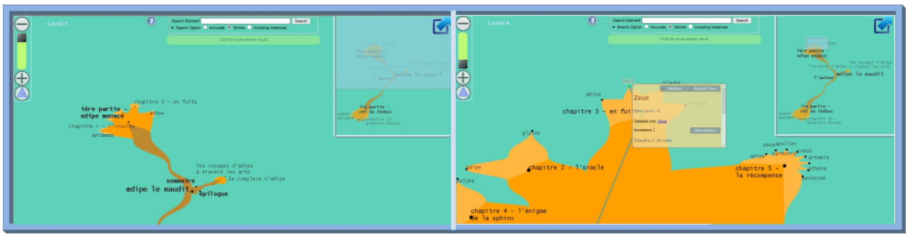

Fig. 2. A children book "Oedipe le maudit" (copyright belongs to SEJER) visualized with the Memory Islands technique. In this visualization, the People identified in each chapter. The zoom and detail-on-demand function provides more in-depth information; the path is added to improve navigation performance; search function allows the young users to search the concepts they think interested.

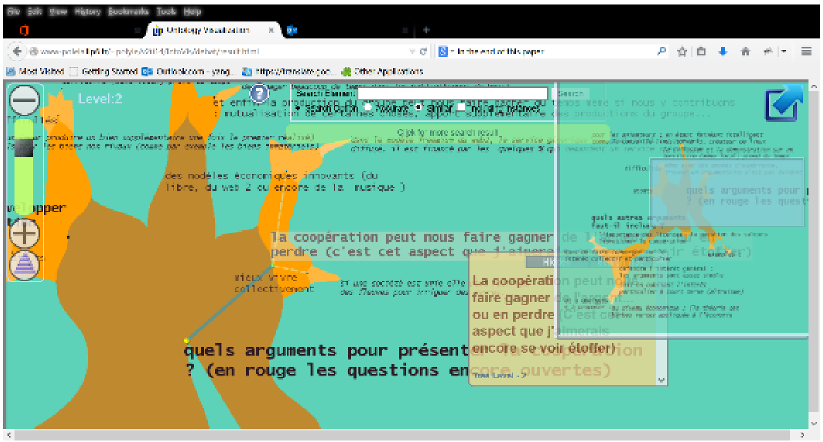

Fig. 3. A public debate in the French Parliament visualized with Memory Islands. Statements are grouped by the type of topic.

3.2 Navigating through the encyclopedia of Philosophy

In this section we will show a knowledge map generated from the ontology - InPho [1], created by the Indiana University to describe the Stanford Encyclopedia of Philosophy (Fig. 4). This ontology doesn't have a structure as simple as a tree (e.g., one node may have multiple parents). The colors show the different levels inside the taxonomy, while the size of each region of the island is proportional to the number of articles in the encyclopedia about the concept represented in that respective region.

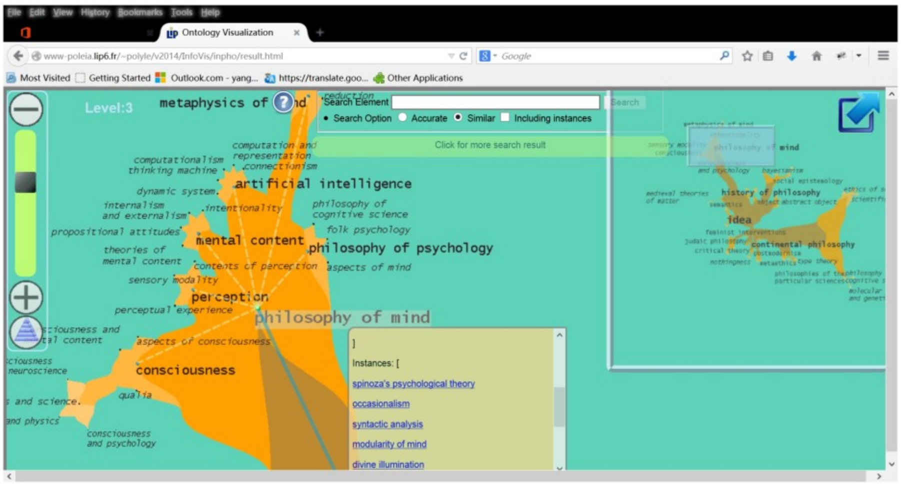

Fig. 4. A knowledge map for the Stanford Encyclopedia of Philosophy. Users can navigate through the Stanford Encyclopedia of Philosophy using Memory Islands and can directly access an article corresponding to a concept in the encyclopedia by clicking the link on the map.

4 Conclusion and Discussion

In this position paper, we have shown some of the knowledge maps that we created by using the Memory Islands. Our technique is taking advantage of the users' familiarity with a tool as common as a map to help them achieve complex visualization tasks. The metaphors we designed enabled users to have in-depth insight into the given knowledge.

We have performed a series of experiments, exposing users to different techniques applied to different ontologies; we found that our technique gives some advantages for the users without experience to navigate and remember ontologies. This is particularly encouraging, since our tool was designed to support the users with their navigation and memorization tasks.

We designed different spatialization algorithms to express the different visualization metaphors, and tested our algorithms with lager ontologies and tree structure. We would like to discuss with Knowledge Mapping experts about the quality of the mappings we could find in the knowledge in order to improve our visualization technique.

Knowledge orders (organization) is an important topic for creating a spatial representation. Some may like more symmetrical spatial representation, while others may like to organize the elements by their meanings – putting those that display highly semantic similarity concepts close together and non-related concepts far apart. For this reason, Memory Islands propose a hierarchical reorganization of the information. To this end we designed two different methods for reorganizing the elements in a tree structure: the first method considers the volume of concept (weight), which tries to reorganize the tree structure to generate a symmetric island (Fig. 5). The second method considers the elements using a measure of semantic similarity. An intuitive way of visualizing the semantic similarity of terms is by grouping together closely related terms and spacing more distantly related ones further apart (Fig. 5).

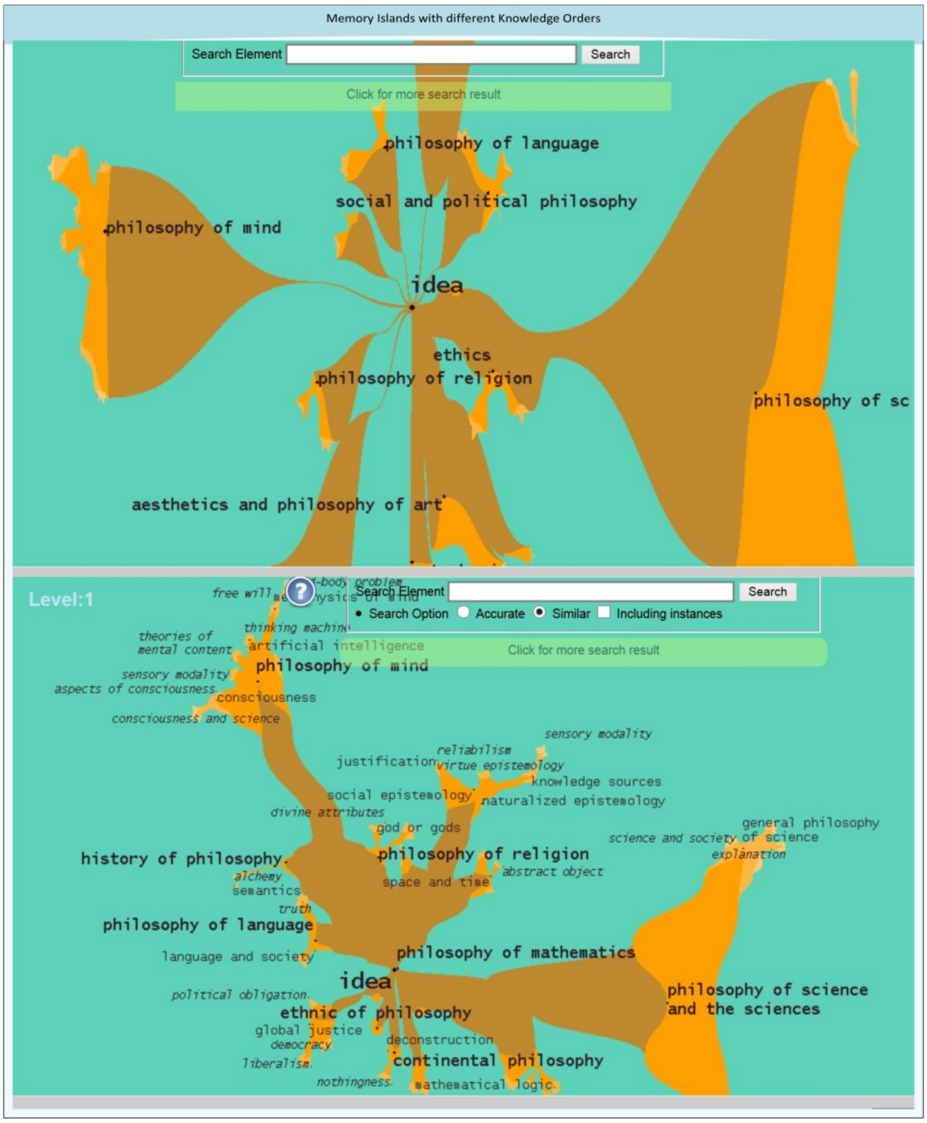

Fig. 5. Two maps generated for the InPho Ontology. The first one display the concepts more

Additionally, we have introduced a small search engine in our visualization tool. Currently performing only simple searches over the taxonomies and the instances, we aim at developing it to function over ontology structures, taking advantage of different relationships inside the ontology (e.g., it could give related results or close results for a query, in addition to the exact matches. We would like to discuss with the Information Retrieval community about the search functions that we should embed in our tool. We think that feedback from the IR community will surely render our resulting knowledge maps more powerful and useful for both users and knowledge designers and engineers.

Acknowledgments

We would like to thank Mihnea Tufis of Lip6 for his valuable comments and suggestions for improving this paper. This work was partially supported by a grant from the French National Research Project (Investissements d'avenir) LOCUPLETO. This work has been done within the Labex OBVIL project and in collaboration with COST Action TD1210 (KNOWeSCAPE).

References

1. Buckner, C., Niepert, M., Allen, C.: InPhO: The Indiana philosophy ontology. APA Newsletters - Newsletter on Philosophy and Computers, 7 (1), pp. 26-28. (2007)
2. Cockburn, A.,: Revisiting 2D vs 3D implications on spatial memory. In Proceedings of the fifth conference on Australasian user interface - Volume 28 (AUIC '04), Cockburn, A. (Ed.), Vol. 28. Australian Computer Society, Inc., Darlinghurst,, Australia, 25-31. (2004)
3. Darken, R., Sibert, J.L.: Navigating large virtual spaces. International Journal of Human-Computer Interaction, 8, 49-72. (1996)
4. Gansner, E., Hu,Y., Kobourov, S.: Chapter: Viewing Abstract Data as Maps, in book: Handbook of Human Centric Visualization, Publisher: Springer, Editors: Weidong Huang, pp.63-92. (2013)
5. Mosallem, Y. Abi-Haidar, A. and Ganascia J.G.: Unsupervised Named Entity Recognition and Disambiguation: An Application to Old French Journals, volume 8557 of the Lecture Notes in Computer Science series. In Proceedings of ICDM 2014.St. Petersburgh, Russia. (2014)
6. Skupin, A., Fabrikant, S. I.: Spatialization methods: A cartographic research agenda for non-geographic information visualization. Cartography and Geographic Information Science, 30(2), 99−120. (2003)
7. Thorndyke, P., and Hayes-Roth, B,: Differences in Spatial Knowledge Acquired from Maps and Navigation. Academic Press. Cognitive Psychology 14, pp 560-589. (1982)
8. Yang, B and Ganascia, J-G,: Memory Islands: an approach to cartographic visualization, In: Classification & visualization: interfaces to knowledge: proceedings of the International UDC Seminar 24-25 October 2013, The Hague, The Netherlands. Edited by Aida Slavic, Almila Akdag Salah, Sylvie Davies. Wurzburg: Ergon Verlag, pp 137-152. (2013)