Thoughts Reflection Machine

Claus Atzenbeck*

Hof University
Institute of Information Systems
Hof, Germany
claus.atzenbeck@iisys.de

Daniel Roßner

Hof University
Institute of Information Systems
Hof, Germany
daniel.rossner@iisys.de

ABSTRACT

This blue sky paper presents the *Thoughts Reflection Machine* (TRM) which combines hypertext technologies and intelligent components. Using hypertext, the TRM provides means to its users to express or communicate their thoughts and ideas. Furthermore, the machine suggests relevant information that trigger users' creative thinking. The TRM is an approach towards a tight cooperation between human and machine supporting both in their specific tasks in which they are most excellent in: creative problem solving respective computation of huge data sets.

CCS CONCEPTS

• **Human-centered computing** → **Hypertext / hypermedia**; *Graphical user interfaces*; *User centered design*; • **Software and its engineering** → *Software infrastructure*.

KEYWORDS

hypertext; thinking; cooperation; AI; augmentation; intellect; intelligence; man-machine; structures; context; infrastructure

ACM Reference Format:

Claus Atzenbeck and Daniel Roßner. 2020. Thoughts Reflection Machine. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media (HT '20)*, July 13–15, 2020, Virtual Event, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3372923.3404837>

1 INTRODUCTION

Hypertext has been described by Conklin as a “computer-based medium for thinking and communication” [8]. Hypertext enables users to express their thoughts using nodes and links. As such, hypertext is a medium that is tightly connected to the user.

There are many pre-web hypertext research publications that focus on user centric issues, e.g. structuring information in KMS [1], representing structures in NoteCards [14], or the representation of links [cf. 24]. In the 1990s the hypertext research community started to discuss several further domains beside the so-called navigational hypertext (i.e. node-link structures), including spatial hypertext

[16], taxonomic hypertext [17] or argumentation supporting structures [9]. The different structure types help users in solving various organizational problems.

The Web has been developed with a similar focus: helping its users to publish information (i.e. websites) interconnected with links (i.e. anchors containing URIs). The next “generation” of the Web supported closer cooperation, e.g. writing collaboratively on wikis or commenting on blogs. Later, social media and social networks brought content sharing to the broad public.

In recent years AI became more and more important in our daily lives. In many cases AI *automates* tasks that otherwise would be performed by humans [3], e.g. autonomously driving vehicles. In the early hypertext days, however, the hypertext pioneer Douglas Engelbart proposed to *augment* human capabilities in contrast to automation [12]. This leads to the discussion about how augmentation of human intellect (which is a fundamental hypertext topic [cf. 6, 8, 10, 23]) can be enhanced by AI.

In this paper we propose a system, the *Thoughts Reflection Machine* (TRM), that fosters human creativity by combining hypertext technologies and intelligent components. The TRM augments human intellect by externalizing the human thinking process and connecting it to the machine's knowledge bases. The discussion starts with presenting the basic idea in Sect. 2, followed by a description of the proposed system (Sect. 3). Further, we will describe two user scenarios in Sect. 4 and conclude the discussion in Sect. 5.

2 BASIC IDEA

The basic idea behind the *Thoughts Reflection Machine* (TRM) is – speaking with Conklin – to provide a “medium for thinking and communication” [8] combined with computer-based intelligent components, including analytics and recommender functionality. The latter should reflect the user's thoughts and foster creative processes that support problem solving. Because of its basic feature set we classify the whole TRM system as a hypertext system.

In [3] we discuss in greater detail the intersecting area of hypertext (for expressing human associations) and machine intelligence (for computing data and recognizing patterns). We argue that this can be seen as a combination of *augmentation* (human using hypertext) and *automation* (machine using AI components).

Along that line one may also think of recommender systems as used, for example, in social media or online shops. The recommendations the user receives are based on previously collected data without giving the user the opportunity to *explicitly* express own thoughts. It is not possible for the user to express something like: “the algorithm book's cover picture I see in my online book store reminds me of flowers, which make me think of asking my friend to help me with my garden work tomorrow.”

*Corresponding author

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
HT '20, July 13–15, 2020, Virtual Event, USA
© 2020 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7098-1/20/07.
<https://doi.org/10.1145/3372923.3404837>

The goal of the TRM is to capture relevant associations popping up in the users' heads. It provides the freedom to users to freely associate information and suggests further relevant information pushing users' creativity. It aims at supporting users in thinking "outside the box."

3 THOUGHTS REFLECTION MACHINE

3.1 Background

TRM permits humans to think aloud while it attempts to suggest relevant information from its computed knowledge. The main interface for visualization and user interaction is a 2D space on which informational units are displayed and put in relationship. As such, the TRM classifies as a *spatial hypertext system*, as discussed in various publications [e.g. 21].

Further below in this section we describe the general workflow of the TRM, depicted in Fig. 1. Parts of that workflow refer to our component-based open hypermedia system (CB-OHS) *Mother*, which is described in detail in previous publications [4, 5]. *Mother*'s architecture consists of three layers:

- (1) *Midgard* ("user layer") includes any user application or visualization component
- (2) *Asgard* ("structure layer") takes care of various hypertext structure types
- (3) *Hel* ("knowledge layer") holds computer generated knowledge bases

It is worth mentioning that *Asgard* includes a spatial structure service that is one of the important components for the TRM. It contains so-called *spatial parsers* [cf. 20] enabling the machine to interpret the relations between objects on the space based on their position or visual appearance.

3.2 Human Thinking

The very start of the TRM's workflow, as indicated on the very left in Fig. 1, is the human user thinking aloud. Thinking aloud lowers the number of interruptions that may come up in cases when the user has to enter ideas into the machine via keyboard or touch input device. Even more: users do not have to pay attention to the machine at all; instead, they may concentrate exclusively on their thoughts while the machine is "taking notes" of what is said.

However, as we will mention below, human speech may not be the only input type for the TRM. Imagine already written texts, e.g. an electronic book that a user reads and for which he/she wishes to see a generated overview map of the content. Other inputs may include photographs on which computer vision technologies detect items or OCR (optical character recognition) detects text. In the future, even advanced brain-computer interfaces (BCI) may be used to visualize concepts derived directly from users' brains.

3.3 Text Analytics and Visualization

Our current focus for the first version of the TRM is on texts derived from users' speech (thinking aloud) or already existing text documents that the user is reading along.

When it comes to *automatic speech recognition* (ASR), the latter scenario is already covered well by recent software [13]. Spontaneous speech on the other hand is harder to handle, because of the

speakers variability in speaking and thinking aloud [22]. In recent years, many proposed ideas [e.g. 11, 15] and more training data lead to significant improvements on word error rates, yet it will be difficult to apply the same NLP tool chain as for written text or transcribed planned speech.

The main purpose of the TRM is to extract keywords from speech, hence it is not necessary to have a complete transcription, as long as the system is fed with those words. Structured input can be computed by traditional part of speech (POS) tagging retrieving the context of keywords in a sentence. When the application domain is known, named entity recognition (NER) may add more helpful metadata. Spontaneously spoken words, even if they are understood completely, do not form a well structured sentence, especially when a speaker is verbalizing thoughts and ideas in a brainstorming scenario. Modeling these inaccuracies into POS tagging and NER helps to improve error rates [7].

In order to visualize extracted words – more precisely: their relations to each other – some sort of associations have to be determined. We use *weights*, ranging from 0 to 1, which encode the strength of such a relations. Weights depend on considered features and can either be retrieved from a knowledge base in *Hel* (passive) or calculated from linguistic, syntactic or temporal features of the speech (active). Actively inferred weights can be refined over time, because entities may appear more than once and in different contexts. Ideally, these weights draw near the original intention of the user and follow the progress of his/her thoughts.

Visualization happens in a 2D space: keywords are laid out such that spatial proximity to each other serves as a visual metaphor for the inferred weights. As described, these weights may get refined, new keywords appear, less important keywords disappear. To handle this dynamic environment we suggest the usage of a spring based metaphor [cf. 19], because approaches based on physics offer a natural behavior of objects when forces are applied, e.g. a new keyword is added. Other visual properties can encode other metadata, e.g. color to represent named entities or size to show the average weight of an entity. As long as the user does not interact with the space (cf. Sect. 3.4), all visual properties are controlled by the system only.

3.4 Context Creation

So far the visualization considers two types of nodes: *user nodes* (i.e. entities derived from the user's thinking process via NLP) and *suggestion nodes* (i.e. relevant information coming from existing *Hel* knowledge bases). The placement of those nodes is completely based on the machine's computation of associations.

The TRM, however, provides means to the user to create *contexts*. As a kind of filtering, such contexts take immediate effects on both the visualization and the computed suggestion nodes. The context is created by anchoring nodes on the 2D space. Those *anchored nodes* keep their positions and visual attributes. They define the context in which any user or suggestion node is placed.

The associations between anchored nodes are computed by the spatial parsers that are part of *Asgard*'s spatial structure service. Given the resulting graph, *Hel* suggestions apply to the whole structure rather than individual nodes. The returned suggestion

Figure 1: Workflow of the proposed Thoughts Reflection Machine (TRM)

nodes as well as the derived user nodes are then placed semantically meaningful within the given context.

Figure 2 provides a demonstration of the basic user interface. The screenshot was taken from [18] and shows a scenario in the movie domain with movie titles, directors, actors etc. Large nodes are objects entered by the user. Smaller nodes are suggestions added by the machine. Those suggestion nodes depend on the user created context and may appear or disappear depending on how the user modifies the structure of the user nodes. The above mentioned entities extracted from the user’s speech are not considered in the current system and, thus, do not appear in the given screenshot.

Even though Fig. 1 suggests a linear workflow, the various parts of thinking aloud, creating context by anchoring nodes, analysis of the text input, the space, querying of suggestion nodes, and fluid, animated visualization are happening in parallel. This enables the user to switch between thinking aloud, taking notes on the 2D space, creating context or getting inspired by the visualization or the machine’s suggestion nodes whenever needed.

3.5 Result of the TRM Process

The TRM lets users and machine populate 2D information spaces with knowledge that changes over time by adding further information to the space or by shaping the context in which computed information should be presented. This fosters human creativity.

The result is a spatial hypertext representing both the process and the outcome of human thinking enhanced by the machine’s capability of computing or searching huge datasets. This knowledge space may then, for example, be further developed at a later point of time, communicated to others, or archived.

4 CASE SCENARIOS

In order to provide a better understanding of the TMR, this section includes descriptions of two scenarios. The first is about TRM supporting spatially separated colleagues brainstorming about solutions to some problems using a virtual conference room. The second is about a student reading a novel, using the TRM to experiment with different contexts, i.e. views on the plot, story line or characters.

4.1 Scenario 1: Team Cooperation

This scenario is about *synchronous cooperation of spatially distributed people*. Imagine Otto and his team working on some solutions to certain problems at hand. Due to a pandemic the team members have to work at home. They schedule a virtual meeting for discussing their work.

During the meeting the participants’ audio is captured by the TRM, including information about the respective speaker. The TRM processes the audio input and its metainformation and presents the most important terms in a window that is shared among all participants. The presentation implies the computed relationships between terms, considering the given metainformation. Metainformation can be any information that makes computing relationships between topics more accurate, e.g. who of the participants mentioned which idea or in which mood the speaker was (angry, aggressive, happy etc.).

The analysis and presentation happens in “real time” and is frequently updated such that any participant can follow the current state of the discussion in the TRM window. Furthermore, the system presents additional terms from the knowledge base and presents those accordingly.

While the discussion and TRM’s speech/text analysis takes place, Otto takes over the role of a moderator and starts anchoring selected ideas on the space. By doing so he creates a context of terms which are most important to the team. He implicitly associates them by organizing them spatially on the canvas, changing their size or applying color or other visual attributes to them.

TRM’s structure layer Asgard computes the relations of all anchored nodes and queries the knowledge base for relevant information. As a result suggestion nodes are displayed that match the *context as a whole* rather than nodes individually. By skimming the space during the discussion, Otto and his team detect some new information of which they have not thought yet and anchor those on the space.

At the end of the session the space shows the suggested ideas mentioned by the participants themselves as well as information suggested by the machine that was added by the participants. The created knowledge on the 2D space has evolved during the discussion with little input from the participants. It represents solutions to the problem the team was working on. Otto and his team can now move one step further and transform the knowledge map into a more formal structure for further processing.

4.2 Scenario 2: Books in Context

This scenario is about *live contextualizing texts* during the reading process. Imagine Sue, a college student studying English literature. She attends a seminar about Tolkien’s epic fantasy novel *The Lord of the Rings*. She is asked to read the books in order to get a better understanding of the various plots and characters.

Sue reads a digital edition of the book that is connected to the TRM. While she turns pages, the system considers each new opened page as input to be processed and displayed on the screen. Thus, Sue can follow how the most important terms, characters etc. in

Figure 2: Application screenshot showing the spatial hypertext with user created nodes (large) and computer-generated suggestions (small); figure taken from [18]

the story evolve while she is reading. In other words, the TRM proposes a view of the story from the beginning to the spot Sue is currently busy reading. Additionally, Sue can speak aloud any idea or association that comes up during her reading. The TRM will capture her spoken words, process them and add the most important ones onto the space. Beside that the system adds further suggested information which may come from other parts of the novel itself as well as from other relevant books, websites, Sue's personal notes or other sources.

Not only can Sue be reminded of the various characters and issues mentioned in the story so far, but she may also start creating a context in which the story should be presented: She starts organizing a few characters and items mentioned in the story on the 2D space. The system computes the relationships between the anchored nodes and proposes further relevant information from the knowledge bases.

For example, Sue is in particular interested on the impact the ring has (directly or indirectly) on the relationship between the wizards Gandalf and Saruman. She moves items representing both wizards close to each other and applies the same color in order to express a strong connection. Additionally, she places an object representing the ring. TRM's structure aware components interpret the space and display relevant information in relation to the given context.

During the reading phase, Sue gets notified by the TRM about some relationships implicitly mentioned in the plot which she would not have discovered otherwise. She adapts the context several times based on her respective focus while reading the story. At the very end she has created a knowledge space on which the most important aspects of the story relevant to her seminar are presented. Furthermore, using TRM's history feature, she may "travel back" in time and review the creation process or suggested nodes [cf. 2], e.g. for preparing upcoming exams.

Summarizing, for the student Sue the TRM combines note taking, entity extraction, recommender system and search facilities. It fosters her creativity and ability to associate, supported by the machine's capability to process or search masses of data/information. Similar to the first scenario described above she can use the result for further processing (e.g. writing an essay about the novel) or discussing it with other students.

5 CONCLUSION AND FUTURE WORK

In this paper we proposed the *Thoughts Reflection Machine* (TRM) that is based on our CB-OHS *Mother* and combines hypertext functionality with data processing and intelligent components for structure interpretation. We have shown in two usage scenarios how the TRM helps individuals or teams in reflecting their ideas and combining them with other's information or computed suggestions. We also described how emerging contexts are used to provide specific views on the given information.

The use of such a system is potentially unlimited, including any setting in which the users need to be creative and have the desire to be supported by the machine in an intelligent way.

Most components needed for the TRM already exist today, including speech to text or NLP technologies or the spatial parsers. However, the combination of those technologies toward the described system is completely new to our knowledge.

The next steps include a detailed description of the technologies needed. From that we plan to integrate missing parts in our system *Mother* in order to reach a working demonstrator. The goal is to use the created TRM demonstrator for user tests. We want to derive insights in the productive use of such a system, including knowledge about how the system can be improved (e.g. components, architecture, technology etc.) or how it can be used in various application domains.

REFERENCES

- [1] Robert M. Akscyn, Donald L. McCracken, and Elise A. Yoder. 1988. KMS: a distributed hypermedia system for managing knowledge in organizations. *Commun. ACM* 31, 7 (1988), 820–835.
- [2] Claus Atzenbeck and David L. Hicks. 2009. Integrating Time Into Spatially Represented Knowledge Structures. In *Proceedings of the International Conference on Information, Process, and Knowledge Management (eKNOW)*. IEEE Computer Society, 34–42.
- [3] Claus Atzenbeck and Peter Nürnberg. 2019. Hypertext as Method. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media (HT '19)*. ACM, 29–38.
- [4] Claus Atzenbeck, Daniel Roßner, and Manolis Tzagarakis. 2018. *Mother – An Integrated Approach to Hypertext Domains*. In *Proceedings of the 29th ACM Conference on Hypertext and Social Media*. ACM Press, 145–149.
- [5] Claus Atzenbeck, Thomas Schedel, Manolis Tzagarakis, Daniel Roßner, and Lucas Mages. 2017. Revisiting Hypertext Infrastructure. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media 28th ACM Conference on Hypertext and Social Media*. ACM Press, 35–44.
- [6] Belinda Barnet. 2018. Boosting Human Capability. In *Proceedings of the 1st Workshop on Human Factors in Hypertext (HUMAN '18)*. ACM, 1.
- [7] Frédéric Béchet, Allen L. Gorin, Jeremy H. Wright, and Dilek Hakkani Tür. 2004. Detecting and extracting named entities from spontaneous speech in a mixed-initiative spoken dialogue context: How May I Help You? *Speech Communication* 42, 2 (2004), 207–225.
- [8] Jeff Conklin. 1987. Hypertext: an introduction and survey. *Computer* 20, 9 (1987), 17–41.
- [9] Jeff Conklin and Michael L. Begeman. 1987. gIBIS: a hypertext tool for team design deliberation. In *Proceedings of the ACM Conference on Hypertext*. ACM Press, 247–251.
- [10] Jeff Conklin, Albert Selvin, Simon Buckingham Shum, and Maarten Sierhuis. 2001. Facilitated hypertext for collective sensemaking: 15 years on from gIBIS. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia*. ACM Press, 123–124.
- [11] Richard Dufour, Vincent Jousse, Yannick Estève, Frédéric Béchet, and Georges Linarès. 2009. Spontaneous Speech Characterization and Detection in Large Audio Database. In *13th International Conference on Speech and Computer (SPECOM 2009)*.
- [12] Douglas C. Engelbart. 1962. *Augmenting Human Intellect: A Conceptual Framework*. Summary Report AFOSR-3233. Stanford Research Institute.
- [13] S. Furui. 2002. Recent progress in spontaneous speech recognition and understanding. In *2002 IEEE Workshop on Multimedia Signal Processing*, 253–258.
- [14] Frank G. Halasz, Thomas P. Moran, and Randall H. Trigg. 1987. NoteCards in a nutshell. In *Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface (CHI '87)*. ACM Press, 45–52.
- [15] H. Hofmann, S. Sakti, R. Isotani, H. Kawai, S. Nakamura, and W. Minker. 2010. Improving spontaneous English ASR using a joint-sequence pronunciation model. In *2010 4th International Universal Communication Symposium*. 58–61.
- [16] Catherine C. Marshall, Frank G. Halasz, Russell A. Rogers, and William C. Janssen. 1991. Aquanet: a hypertext tool to hold your knowledge in place. In *Proceedings of the 3rd ACM Conference on Hypertext*. ACM, 261–275.
- [17] H. Van Dyke Parunak. 1991. Don't link me in: Set based hypermedia for taxonomic reasoning. In *Proceedings of the 3rd Annual ACM Conference on Hypertext*. ACM Press, 233–242.
- [18] Susanne Purucker, Claus Atzenbeck, and Daniel Roßner. 2019. Intelligent Hypertext for Video Selection: A Design Approach. In *Proceedings of the 2nd International Workshop on Human Factors in Hypertext (HUMAN'19)*. ACM, 19–26.
- [19] Daniel Roßner, Claus Atzenbeck, and Tom Gross. 2019. Visualization of the Relevance: Using Physics Simulations for Encoding Context. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media (HT '19)*. ACM, 67–76.
- [20] Thomas Schedel and Claus Atzenbeck. 2016. Spatio-Temporal Parsing in Spatial Hypermedia. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media*. ACM, 149–157.
- [21] Frank M. Shipman, J. Michael Moore, Preetam Maloor, Haowei Hsieh, and Raghu Akkapeddi. 2002. Semantics happen: Knowledge building in spatial hypertext. In *Proceedings of the 13th Conference on Hypertext and Hypermedia*. ACM Press, 25–34.
- [22] Henrik Schulz and José Adrián Rodríguez Fonollosa. 2013. Modelling the effects of spontaneous speech in speech recognition. In *Proceedings of the Speech Processing Conference*.
- [23] Andries van Dam. 1988. Hypertext '87: Keynote Address. *Communication of the ACM* 31, 7 (July 1988), 887–895.
- [24] Harald Weinreich, Hartmut Obendorf, and Winfried Lamersdorf. 2001. The look of the link – concepts for the user interface of extended hyperlinks. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia*. ACM Press, 19–28.