



University of HUDDERSFIELD

University of Huddersfield Repository

Wang, Jing, Ho, Yufang, Xu, Zhijie, McIntyre, Dan and Lugea, Jane

A Visualization Method for Understanding Forensic Statements

Original Citation

Wang, Jing, Ho, Yufang, Xu, Zhijie, McIntyre, Dan and Lugea, Jane (2016) A Visualization Method for Understanding Forensic Statements. In: IEEE Information Visualization (InfoVis), 23rd-28th October 2016, Baltimore, Maryland, USA. (Unpublished)

This version is available at <http://eprints.hud.ac.uk/id/eprint/29990/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>

A Visualization Method for Understanding Forensic Statements

Jing Wang*

Yufang Ho†

Zhijie Xu‡

Dan McIntyre§

Jane Lugea¶

University of Huddersfield
Huddersfield, West Yorkshire, UK, HD1 3DH

ABSTRACT

Forensic statements are lengthy and contain large amounts of complex information. Consequently, it is often difficult for readers of such reports to identify connections between disparate pieces of evidence and to properly and objectively assess their value to the case in question. Readers have no alternative but to rely on intuition and experience to make sense of the complex arguments and propositions arising from forensic evidence. This research investigates the opportunities in the convergence of linguistic approaches to extracting and reconstructing the cognitive structure, i.e. “Text-Worlds”, in a statement, and the computerized operational settings for enabling effective and hopefully more accurate interpretation of forensic discourse through visualization. This will be of benefit to a wide range of stakeholders, including investigating officers, prosecuting and defence counsels, judges and jurors.

Index Terms: H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing—Linguistic processing;

1 INTRODUCTION

Forensic statements are the text records made by victims, witnesses and suspects. Those statements, along with other physical evidence, are often used to reconstruct truthful crime stories during criminal investigations and court proceedings. However, legal practitioners usually have to face necessarily detailed, extremely complicated and large amounts of text records. It is often challenging for them to identify connections between disparate pieces of linguistic evidence, and to properly assess their probative values to the case in question. Various approaches have been developed for evidence representation, manipulation and automated analysis. However, a rigorous method that can facilitate the interpretation of detailed forensic statements and assist in objective judgments is still under-developed. Very little research attempt has been made in systematically representing cognitive linguistic structures in a statement into an intuitive visualization format for forensic applications.

The research has built on and improved existing techniques in the field by operationalizing a sophisticated cognitive model of human discourse processing known as Text World Theory (TWT) [1]. A linguistic information visualization system named as Worldbuilder has been developed in this research. Worldbuilder is a web-based visual inference system that is designed to help users to understand complicated forensic statements. The system applies a mark-up scheme that encompasses the range of descriptive categories found in TWT to raw statements annotation. The mark-up scheme provides an objective and replicable structure that could be useful in examining which elements of the text are likely to be important for criminal investigation. An interactive visualization platform is then

deployed for the purpose of assisting users in understanding the crime stories reconstructed from the annotated data.

This poster highlights two contributions: firstly, a TWT mark-up scheme has been developed, which provides a structured framework for quantitative analysis of complicated information from human language and raw text; secondly, cognitive structures of forensic statements can be visualized through a graphical human computer interface, which helps forensic linguistic experts or potential end-users (e.g. investigators, judges, prosecuting and defence counsels, etc.) to understand and evaluate those forensic statements.

2 THE WORLDBUILDER ANNOTATION AND VISUALIZATION SYSTEM

The theory used in this research, named as TWT, is situated within the tradition of cognitive stylistics, focusing on relating linguistic choices to cognitive structures and processes that underlie the production and reception of language. In our research, TWT has been operationalized as a structured framework based on the nature of the forensic statements as shown in Figure 1. The cognitive model of human discourse processing divides all discourse situations into three manageable levels of conceptual activity: discourse world, text worlds, and world switches. In this framework, the worlds are constructed via world building elements and linked by the world switches, e.g. a shift in time and/or location from one world to another. These elements constitute the fundamental quantitative inputs for the visualization process. Detailed discussions on the TWT is beyond the scope this poster. Interested readers are suggested to refer [1] for more information.

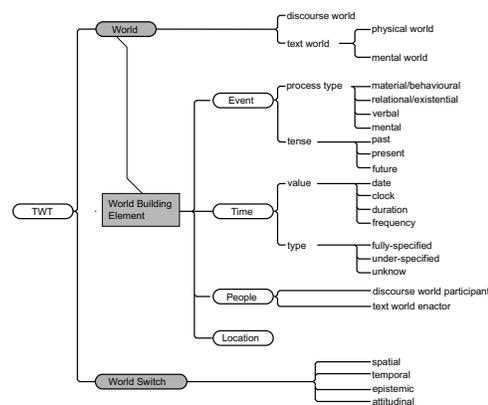


Figure 1: TWT cognitive framework.

Based on the cognitive framework, Worldbuilder has then been developed with two primary objectives: (a) to assist human researchers in the manual annotation of the key linguistic elements associated with the cognitive structure in a forensic text; and (b) as a result of the mark-up of data, to improve computer-based techniques for visualizing complex information and accordingly enhancing the readability of forensic statements. This mark-up scheme is a significant improvement on existing practices due to its systematic and replicable operational approach. The research

*e-mail: jing.wang@hud.ac.uk

†e-mail: y.ho@hud.ac.uk

‡e-mail: z.xu@hud.ac.uk

§e-mail: d.mcintyre@hud.ac.uk

¶e-mail: j.lugea@hud.ac.uk

work is focused on the development of the Worldbuilder annotation interface and the resulting visualization outputs.

To present the text-worlds cognitive structure effectively, both aesthetic form and functionality need to go hand in hand, providing insights into complex data by communicating its key aspects in an intuitive way. Worldbuilder contains two major consecutive operations. At first, raw forensic statements are manually annotated according to the principles and rules of the devised TWT framework (Section 2.1). Then, the digitized and structured data can be interpreted and analyzed through the proposed interactive visualization scheme (Section 2.2).

2.1 TWT Mark-up

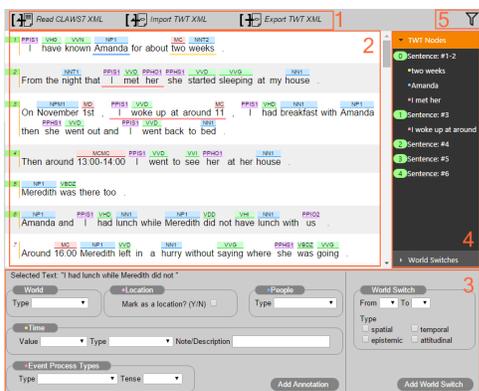


Figure 2: Snapshots of UI

As shown in Figure 2, Worldbuilder is a web-based system developed in this research for assisting forensic linguistic analysis. The system is developed by using the Hypertext Mark-up Language (HTML5) and JavaScript, which can be readily processed and displayed by mainstream Web Browsers such as Internet Explorer, Safari, Firefox, and Chrome. This approach also provides great portability for the system usage in various legal proceedings.

It is worth noting that the color features at the linguistic annotation stage is driven by the real demands from linguistic analysts for improving user experience. Visualizing raw texts in this manner, especially through highlighting their grammatical properties, can help annotators to locate “world building elements” and “events” more efficiently. The outcome of this annotation process is a diagram of text-worlds projected in each statement (see Section 2.2). The TWT mark-up process in this design will then convert raw statements into structured format for ease of processing data quantification and information visualization.

As shown in Figure 3, the annotation interface contains 5 zones (referred as Z1, Z2, Z3, Z4 and Z5). Z1 is used for project and file management with functions like saving and loading TWT mark-ups as XML files. Z2 provides an editing window with Z3 offering the TWT mark-up tools. The annotation will be automatically highlighted by colored underlines in Z2. Z4 displays the marked world-building elements (time/location/people) and the “events” in a world in a list format. Z5 is designed for user interface visual filtering purposes, that is, for the annotators to decide which grammatical features to be shown in the raw text, to keep the information clean and clear.

2.2 Visualizing TW cognitive diagrams

The TWT cognitive framework is employed for analyzing forensic statements via reconstructing the text-worlds projected in a text. The world building elements, together with world-switches, offer

the building blocks for answering the questions such as WHO-WHEN-WHERE and WHAT HAPPENED around a case. A cognitive diagram is the graphical representation of those mental projections which outline a story presented by a person who is involved in a crime (e.g. victim, witness or suspect) through the formation of the world-building elements (nodes) and the connecting switches. Instead of reading original statements, the potential end-users can easily follow the diagram, identify links and contradictions among elements, and interact with other lay-users through visual deliberation.

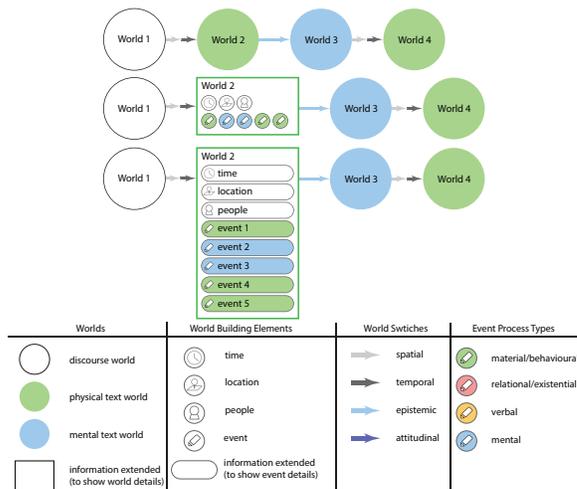


Figure 3: Layout of TW cognitive diagram

As shown in Figure 3, the TW cognitive diagram is designed as a directed network that is capable of storing multiple layers of information. Interactions and animations can also be introduced through augmented transitional operations.

The cognitive diagram is ideal for visualizing the topological structure of text-worlds. As shown in Figure 3, the colored circles represent specific types of worlds (i.e. white indicates discourse world, green indicates a text-world that projects the actual happenings in physical world, and blue indicates a text-world that projects the happenings within ones mind, i.e. mental world). The directed and colored lines indicate different types of world switches.

This initial setup creates a clear and simple view of the entire cognitive structure. Users can gain a brief overview and understanding of the key relationships among all conceptual activities. Interactions and animations are used when exploring the diagram. As illustrated in the middle and bottom of Figure 3, a world node can be expanded to different levels of detail to show its world building elements such as time, location, enactors, as well as event-related information. In addition, the default graph layout is sorted by the world switches, but the diagram can be also sorted by time (i.e. time-line of the crime story), people (i.e. witness and suspects relationships) and location, which can assist objective judgments made by end-users. Another important advantage of using such structured information and representation lies in its potential in comparing different statements, i.e. stories reported by various parties. Worldbuilder is designed to show multiple cognitive diagrams drawn from multiple statements, to support cross-referencing and validation. This function is pivotal when evaluating multiple statements either from repeated interviews of the same subject or from different interested parties.

REFERENCES

- [1] Gavins, J., Text World Theory: An Introduction, 2007: Edinburgh: Edinburgh University Press.