Connecting the Dots:
Representing Real-World Systems Across Multiple Heterogenous Models

Beth Cardier[1], John Shull[2], Alex Nieslen[3], Matthew Garcia[4], Laurie L. Wellman[5], Hargsoon Yoon[6], Richard A. Britten[7], Larry Sanford[8].

[1, 5, 7, 8]Eastern Virginia Medical School, [2, 3]Old Dominion University, [4]Brandeis University, [6]Norfolk State University.
1 bethcardier@sirius-beta.com, 2 jshull@odu.edu, 3 anielsen@odu.edu, 4 mattuesque@gmail.com, 5 wellmall@evms.edu, 6 hyoon@nsu.edu, 7 brittera@evms.edu, 8 sanforld@evms.edu.

Abstract

Real-world events generate information that cannot be captured in a single data set. In fields that grapple with the outcomes of such events, such as medicine and defense, it is desirable to have a modeling paradigm in which multiple media and modeling methods can be logically integrated, ideally with a visual grammar. The problem of linking information across diverse domains is central to effective communication in autonomous human-machine teams, where information must be accurately transferred among different vocabularies, contexts and perspectives.

However, combining multiple heterogeneous models into a unified view is challenging. Barriers include ontological interoperability, in which the fundamental meanings of each component model need to be directly compatible with each other, and representational coherence. We are developing a modeling platform, *Wunderkammer*, in which video, text, image and data from different ontologies can be presented in parallel and visually connected. In our first example, data and expert knowledge from post-traumatic-stress-disorder research in the Sanford-Wellman Laboratory is represented in a single modeling space. This paper describes our work-in-progress towards conveying information from multiple heterogeneous models simultaneously. It is a step towards the larger goal of building an intelligent system capable of automatically integrating multiple heterogenous contexts.

Introduction

By early 2001, US intelligence agencies knew that a small group of Saudi nationals had taken flight lessons in Florida. A different agency had been tracking the evolution of rumors following the crash of Egyptair flight 990, October 31, 1999. Those rumors had been sparked by a suspicion that the National Transportation Safety Board, a US investigative agency, were mistaken in their belief that the Egyptian pilot of this doomed flight had driven his plane into the ocean as an act of suicide. An alternate story had emerged in the middle eastern streets, in which the pilot had heroically saved thousands of lives by plunging the aircraft into the sea instead of allowing it to be flown into Mecca by Mossad, Israel's Intelligence agency. This rumor incubated a conceptual structure which would later seem unforeseen when it became a reality. In this meme, a hijacked plane is crashed into a significant building as an act of revenge.

The 9/11 terrorist attack against the United States in 2001 became a well-known example of how the distribution of information in unrelated forms across many agencies can lose important intelligence. It was framed as a failure to 'connect the dots'. As a consequence, research into the automated integration of multiple contexts – ontological, representational and formal – received a boost, including funding towards this research.

This paper presents a first example from a working prototype modeling platform, *Wunderkammer*. The tool is developed by a team based at the Virginia Modeling, Simulation and Analytics Center (VMASC) at Old Dominion University (ODU), and the experimental data is supplied by a team in the Center for Integrative Neuroscience and Inflammatory Diseases (CINID) at Eastern Virginia Medical School (EVMS). In this tool, multiple, heterogenous models can be simultaneously represented and connected. The tool can accommodate formats of parametric data, video, images, text and semantic information.

Our example case comes from neurobiology, which is another field in which complex real-world information is distributed across many ontological formats and modeling methods. Data collected to measure the impact of stress on emotional and physical resilience can take the form of EEG and local field potentials (LFPs), gene expression data, and biological stress indicators such as sleep loss, temperature and heartrate as well as stress system indices (e.g., cortisol). In addition, there are the semantic descriptions of trauma and its symptoms from the therapist and, in human instances, the patient. Each of these formats represents a different *context*, in the sense that the metrics and presentation of each format are not formally equivalent to the others. This paper gives an overview of how we propose to connect at least three of these models in a single space.

An important feature is the ability to represent these models in parallel, so that new connections can be drawn across the boundaries between them. We have developed a new taxonomy to cross those boundaries based on narrative devices, which are described in other work (Cardier, 2013; 2015) and implemented here. A key feature is the representation of heterogenous information in separate *contexts*, which we refer to as *containers* in the visualization, and *situations* in the theoretical foundations.

Our model features a new way to annotate connections among heterogenous contexts by animating semantic webs among containers over a series of frames. We can also connect this to non-formal semantic information. This differs from related work in semantic visualization such as Discursis and Lexomancer (Angus et al., 2013), in which 'containers' display words or phrases but the nodes which present them do not represent new contexts with new ontological structures. As a consequence, those links are only connections through time. The aim of this tool is to record more "between-the-dots" or hidden expert knowledge than is otherwise possible. A long-term goal is to train an intelligent system based on these semantic structures that can automatically integrate heterogenous models. A rationale for this approach is presented in this work-in-progress.

Defining and Representing Context

Traditionally, computational conceptualizations of *context* have been informed by the capacities of formal representation, such as the work of Minksy (Jahn, 1997) and Schank (1990). In this work, *context* is a closed subset of reality, with a limited scope and well-defined entities. However, if two of these formal representations of context use different structures – *heterogeneous* – they cannot be easily connected or represented together. This problem is well-established in ontological interoperability (Walls et al., 2014) and biological multiscale modeling (Noble, 2015). In both cases, the inability to track real-world phenomena across numerous representational boundaries is a barrier to understanding how real-world systems emerge or become disordered.

At the other end of the representational spectrum is *lived context* and the human perception of it. This notion of context is anchored in aggregations of real-world elements. The aggregation is likely to include physical, social and conceptual factors, and recording these can occur across numerous modeling formats. In this kind of *context*, reality's swim of interdependent agencies constantly changes the role of everything depicted, an evolution that can be anchored in unfolding time. A *context* of this kind will likely contain aspects that only occur in unique circumstances. Devlin describes the challenges of representing a real-world context as trying to capture a horizon that keeps moving (Devlin, 2009). We represent and formalize this kind of context.

In fact, this elusive conceptualization of *context* pervades the humanities, especially from cultural theory, in which numerous contexts affect social identity in the notion of *intersectionality*. Literature uses sophisticated techniques to represent the tensions produced by contextual interaction. We drew on both these notions to develop our taxonomy. A key feature is the capacity to represent change through time; another is the ability to show influence among each contextual frame of reference. This has applications in multiscale sciences; for example, in post-traumatic-stress-disorder (PTSD), a traumatic memory overrides normal biological responses to non-traumatic stimuli.

Our definition of *context* draws from Einhorn and Hogarth's definition of causal attribution, in which meaning occurs in relation to a background frame which provides the terms on which identity is understood (Einhorn and Hogarth, 1986). The 'frame' aspect of context is also echoed in Devlin's logical system 'Layered Formalism and Zooming' (Devlin, 2009), where facts can be nested within other facts. Our research extends his philosophical and logical foundations with a visual representation of *containers.* These frames can change with time, scale and perspective.

Our visualized containers can also be nested, as Devlin's work prescribes. Our representation of this is shown in Figure 1. Containers are a lightweight and intuitive 'drill-down' tool which allows a user to select objects that are by definition 'contained' and also remove them from connection when needed. The containment feature includes the bands which extend indefinitely as long bands (see Figure 2). These bands are sometimes known as a 'swim lanes' in business process models. In our method, swim lanes are also contexts and are represented as open-ended containers.

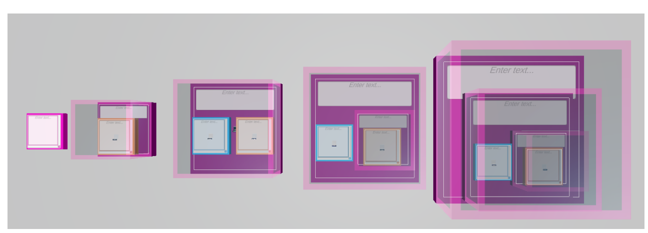

Figure 1: Contexts can be nested in the form of visual containers.

Another feature is the ability to turn any node into any 3D formatted object, including video. This feature enables any object to be easily linked with any other object, regardless of modelling method or media. The user can discretely and uniformly select points around the object via a uniform sphere with points every 30 degrees from center all the way around. This feature results in approximately 96 connection points equally spaced around a sphere. Later, those points will be readable as coordinates by a back-end system. This capability allows the easy linking among objects, allowing the capture of knowledge from the expert researcher, along with the structure to later inform an intelligent back end.

Using these devices, we represent the aspects of context which are novel and changeable, as well as those which are general and repeatable. We are particularly interested in changes of time, scale or subjective perspective as they are important to interpretation of real-world instances. In all of these cases, a core problem is how to represent the transformation of information networks, both visually and formally. Prior work (Cardier, 2014; 2015) describes how we use *narrative* structures to demonstrate these transformations.

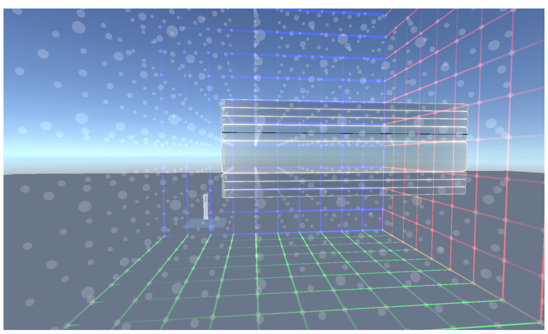

Figure 2: Swim lanes and a grid superimposed over the top. The grid aids alignment of entities across numerous frames.

Narrative structure carries its audience from one state to the next, whether physical, social or ontological (Cardier 2015). It is driven by an incomplete view, in the sense that when new information stops being added, the story is over. Previous connections between story structure and formal models of cognition have already been established (Herman, 2006; Bruner, 1986) yet our take is still unusual. Clarifying our definition of *narrative* is thus worthwhile.

Defining Narrative

Previous papers on this modeling tool have begun with a definition of narrative that follows that of Prince, who observes that a story as an event that causes a change of state (Prince, 1973). This is a good starting place, as the mechanisms of state change are central to our method. However, using a narratological definition can be misleading, as this research is not narratological in its disciplinary scope. Our conceptualization of *narrative* more clearly relates to cognitive science and logic because our goal is to determine the structures which can supplement formal reasoning. For this purpose, we consider that *logic* is a semantic information system in which the mathematical rules of its structures are coherent, consistent and sufficient. By contrast, *narrative* is an information system that tracks the means by which those rules are broken and adjusted to form new coherences. That rule-breaking is not strictly logical; it is something else. Our taxonomy provides an outline of this alternate identity: narrative is a transformation machine designed to intergrate contexts as a means of changing states. Analogical links are an important mechanical aspect of those transitions.

With this in mind, our notion of *narrative* uses and extends definitions with roots in both cognition and computation, such as those of Herman (2006) and Bruner (1986). Both emphasize its transformative, extra-logical qualities. Narratological definitions tend to focus on the relationship between character, event and plot; we anatomize narrative differently, to leverage its role as a system for reasoning-through-change. The structures we extract are those which enable incremental adjustment across heterogenous situations. We frame these operations in terms of computational models of reasoning.

We thus draw several aspects of narrative into our method. One is the basic use of semantics to record non-formal information. Another is the representation of implicit processes of transition, such as the transfer of information among contexts. This generates incremental change , and a third aspect is the generation of new conceptual clusters. We now apply these properties to the representation of a phenomenon that spans many different formal models: post-traumatic-stress-disorder (PTSD).

PTSD example

Our example concerns PTSD and the shifts in neurological circuit activity that occur during its different stages – trauma, learning and recovery. It is chosen because it entails multiple formal and informal measurement scales and representations. In PTSD, the interaction among different bodily and psychological systems is critical to understanding a patient's disorder and recovery. Yet it is difficult to represent these different processes in the same modeling space.

For instance, when a subject develops a fear response to a stimulus, the effect can be measured in multiple ways. We select at least two from a range of concurrent multi-site recordings: EEG and LFP signals (see Figure 3), gene expression data, and biological stress indicators (sleep loss, temperature, and heartrate). These will be modeled beside each other and connected to an immersive 4D model of a brain. To this, we add semantic descriptive observations about the traumatic event and the subsequent response by the researcher. Each of these methods will occupy their own 'swim lanes' (continuous containers). Finally, we annotate the connective structure across the boundaries to link information from one model to another. For instance, measurements during a traumatic event (shock treatment) will be correlated with the equivalent moment in the brain's model, as well as descriptive semantics.

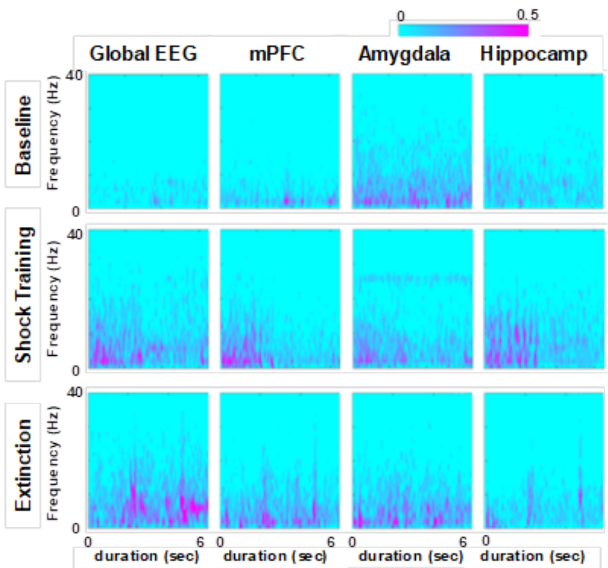

Figure 3: Spectrograms showing differences in activity in the global EEG, and LFPs in the medial prefrontal cortex (mPFC), amygdala and hippocampus during during baseline, and after shock training and fear extinction in a footshock model of traumatic stress.

The new capabilities in *Wunderkammer* allow us to combine multiple models in a single representational space. We have introduced three innovations to achieve this: 1) the technical ability to connect different formal methods and media in the same modeling space, 2) connecting those to an immersive 4D model of a brain, and 3) building a semantic model to record non-formal information alongside these other modelling formats. Currently, a fully integrated model of PTSD only occurs in the mind of a researcher. Our aim is to make more of the expert neurologist's reasoning processes explicit in an immersive modeling environment. Due to the need to include 3D and video artifacts in the space, we use a Unity modeling environment as the basis of the platform. An overview will be discussed and shown in a demo.

Conclusion

Combining multiple heterogenous models in the same space is a step towards a fuller transfer of information across contexts, such as that required by human-machine teams. We produce a first example in our prototype modeling platform, *WunderKammer*. Example data is drawn from a PTSD example supplied by the Sanford-Wellman Laboratory in the CINID at EVMS. Each data form exists in a different discipline and representational system. We connect these into an integrated model to produce a broader picture of cause and effect across experience and neural changes.

References

Angus, D., Rintel, S., Wiles, Janet, 2013. Making sense of big text: a visual-first approach for analysing text data using Leximancer and Discursis, International Journal of Social Research Methodology. May, Vol. 16 Issue 3, p261-267

Bruner, J. 1986, Actual Minds, Possible Worlds. Harvard University Press, Cambridge, MA.

Cardier, B., 2014. Narrative Causal Impetus: Governance through Situa- tion Shift in Game of Thrones, Intelligent Narrative Technologies 7, Association for the Advancement of Artificial Intelligence, eds. J. Zhu, I. Horswill and N. Wardrip-Fruin. AAAI Press: Palo Alto, California, 2-8

Cardier, B., 2015. The Evolution of Interpretive Contexts in Stories, Sixth International Workshop on Computational Models of Narrative, eds. M. Finlayson, B. Miller, A. Lieto and R. Ronfard, Cognitive Systems Foundation, OASICS, V. 45.

Devlin, K., 2009. Modeling real reasoning. In: Sommaruga, Giovanni (Ed.), Formal Theories of Information: from Shannon to Semantic Information Theory and General Concepts of Information. Springer, pp. 234e252.

Einhorn, H. & Hogarth, R., 1986. Judging Probable Cause. Psychological Bulletin, 99, 3-19.

Herman, D., 2006. Genette meets Vygotsky: narrative embedding and distributed intelligence. Language and Literature 15 (4), 357e380.

Jahn, M, 1997. Frames, Preferences, and the Reading of Third-Person Narratives: Towards a Cognitive Narratology, Poetics Today, Duke University Press, 18 (4), Winter: p. 441-468.

Noble, D., 2015. Multi-bio and multi-scale systems biology. Progress in Biophysics and Molecular Biology, 117, 1-3.

Prince, G., 1973. A Grammar of Stories: An Introduction. Mouton: The Hague.

Schank, R., 1990. Tell Me a Story: A new look at real and artificial memory, Charles Scribner's Sons, NY.

Walls, R. L., Deck, J., Guralnick, R., Baskauf, S., Beaman, R., Blum, S., Wooley, J. 2014. Semantics in support of biodiversity knowledge discovery. PLoS One, 9(3), e89606.