

# Graphically Speaking

Editor: Miguel Encarnação

## Isn't It About Time?

Olga A. Kuchar,  
Thomas J.  
Hoeft,  
Susan Havre,  
and Kenneth A.  
Perrine  
*Pacific  
Northwest  
National  
Laboratory*

The simple example in the sidebar illustrates the complexity of interpreting time into common temporal references. To understand all the information presented in this typical news item, an analyst or problem solver might need to reread the paragraph many times to cognitively make links and inferences between pieces of information. Some might go so far as to sketch the events onto a timeline (if you are like us). Interpreting the textual presentation of this detailed information is the bottleneck in our cognitive understanding of the unfolding events. Alternatively, an interactive, visual presentation of this information could increase the information bandwidth and decrease the cognitive load. The timeline sketchers among us understand this intuitively and from experience.

This example is simple when compared to information gathering occurring in many disciplines where information must be extracted from massive amounts

of data, and then linked and structured to distill results. However complex, this critically important activity could inform decisions affecting millions of lives: for example, in finding cures for cancer or preventing terrorist attacks. Even though such information gathering takes place in many domains, this article centers on intelligence analysis of linguistic data. Many intelligence analysts often have to wade through large numbers of articles. Consider the difficulty of creating our sample news story from information embedded in thousands of documents. Say each sentence contains information distilled from 100 documents—analysts would need to find each of these sentences and then link them together to formulate the story. Many different issues arise. For example, knowing which key facts to extract is problematic given the vast amounts of unstructured textual data. Providing analysts with the capability to quickly and accurately understand the inter- and intradocument

temporal relations and information would have tremendous beneficial ramifications for many domains. In this article, we focus on visualizing and interacting with information from such corpora. One of the key attributes in the analysis of event-driven information is time. The purpose of this article is to draw awareness to the time dimension as a unique entity that needs better understanding and attention.

Time plays a key role in all aspects of the intelligence analysis process, from data ingest through analysis methods to the cognitive processes that create intelligence products. However, the concept of time is difficult to grasp and not yet fully understood. Today's analytic tools fall short of satisfying analysts' needs. These tools don't significantly facilitate an analyst's cognitive process when assembling the pieces of a puzzle like those in the previous example because the implicit representa-

### Simple Example

Wednesday 28 April 1999; Posted: 11:33 p.m. EDT (03:33 GMT): Another robbery occurred in southwestern Ontario today, making this the fourth robbery in the past few months. Delaware Bank in Brantford was robbed by three masked individuals who stole \$150,000 in currency and several unknown items from the bank's vault. The bank robbery occurred at 2:30, lasting all of five minutes and injuring eight people. All injured parties were taken to the local hospital where one died on arrival. Two people were released and the remaining people are in intensive care. This robbery is similar to a crime spree that started on the Chinese New Year. The first robbery occurred in the morning at Allegiant Bank in Richmond Hill, with the robbers taking more than \$100,000 in currency. The second robbery occurred about two weeks later at Banner Bank in Ajax and was caught on tape. The robbers arrived just as the bank opened, riding in a white van and wearing black ski masks and black outfits, and carrying automatic machine guns. During the first three minutes, the robbers instructed all of the patrons to face the wall and place their hands on their heads. While two of the robbers watched the patrons, the other robber took the bank manager and instructed him to open the vault. Other video captured the movement in the vault. The vault was opened about five minutes after the robbers arrived. The safe was blown two minutes later, and the robber removed only two safety deposit boxes and placed them in a bag. He then continued to club the manager and was back upstairs, yelling instructions to his buddies as they left the bank, not more than 20 minutes after they arrived. What was unusual was that no alarms were sounded. The third robbery happened at night at Carter Bank in Brampton, with only nearby homeowners mentioning that they do not remember hearing the bank alarm, only dogs barking for a while....

tion of time found in these tools and their visualizations curtail analysts' opportunities to exploit their data's time dimension. There is a need to revisit and understand the nature of the role time plays. In this article, we briefly discuss why time is difficult to grasp, the need for visualizations and interactions to deal with time, and some thoughts about a temporal analytic discourse. Overall, time cannot be left as an afterthought when developing visualizations. We should consider time as a first-class object in its own right, rather than an implicit attribute, and we must make it as interactively manipulable as any other object.

### Why is time difficult?

Although time has been studied by philosophers and scientists for more than 2,500 years and humans have evolved in their understanding, description, and measurement of time,<sup>1</sup> many questions about time remain:<sup>2</sup>

- what time actually is;
- how to work with the metaphor of time's flow;
- whether tense or tenseless concepts are semantically basic; and
- what is the proper formalism or logic that captures the special role that time plays in reasoning.

Humans and their environment are inherently temporal in nature as demonstrated when we try to piece together intelligence evidence. By temporal analytics, we mean that aspect of the knowledge discovery process focused on actively exploiting time to identify patterns and temporal relationships in data.

Time is an important variable. People consider time as an independent quantity since our perception is that we have no control over it. Time is an ever-present thread that can potentially help tie events together. To get this constant thread, a reference frame must underlie the events to create order to evidence and to enable reasoning about it. A reference frame is a standard point of view or perspective for making observations and judgments.<sup>3</sup> Scientists throughout history have created such reference frames and deemed them as absolute time, such as the Gregorian calendar or universal time (see <http://www.worldhistory.com/wiki/T/Time.htm>). For an intelligence analyst, absolute time is the easiest to deal with, letting an analyst create a known reference frame to capture events. There are difficulties, however, with incomplete references to absolute time. For example, "On May 3rd, the USS Defiance left Seattle toward Cape Hope"—the year is unspecified. Another example: "The bank robbery occurred at 2:30 ..."—the date and time zone are unspecified. These examples show how even absolute time might not be semantically fully resolved and might pose difficulties when dealing with global information exchange.

The second example also introduces the need for relative time, a greater challenge to analysts. Humans are good at describing events in a relative manner, using

words such as today, second, yesterday, fortnight, before, after, and Tuesday. The problem occurs when an analyst must normalize these different reference systems into a common framework, with the added difficulty of not knowing which reference system a person might be using. In these circumstances, if an analyst cannot map to a common reference system, the data remains meaningless. For example, "This robbery is similar to a crime spree that started on the Chinese New Year ..."—if an analyst cannot map the Chinese New Year to an actual date on the calendar (in this case, 16 February 1999), then he or she can't use this information as it doesn't help put other information into a reference framework. When normalizing temporal information into a common framework, difficulties often arise around such issues as ambiguity, conflict, uncertainty, biases, causality, scalability, data half-life, and missing or incomplete data.

### Time and analysis

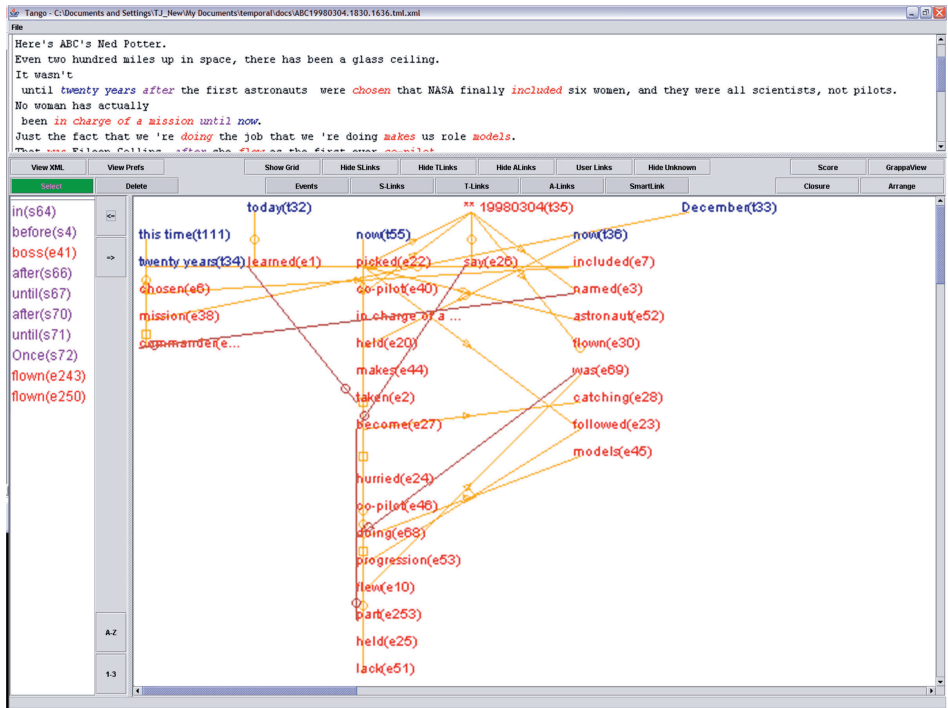
Time is a difficult and ambiguous topic, as evidenced by the variety of representations and metaphors humans have developed to discuss this intangible dimension. When dealing with time, analysts must

- disambiguate it to compare events and durations,
- clearly understand their perceptions of time in a data corpus, and
- have a methodology to build a clear analytic judgment including temporal information.

These three issues start forming some of the requirements for a temporal analytic discourse (TAD) that we define as technology-mediated dialogue between analysts and their temporal information. TAD supports the development and expression of a situation based on time events contained within the data and their relationships contained explicitly or implicitly in data. Due to the massive data that analysts must process, there is a need for visual analytic tools to support such a discourse.

Many of you will argue that the technology does not exist that can truly extract all the information from text. We would agree. However, the automatic recognition of temporal and event expressions in natural language text has recently become an area of intense research in computational linguistics and AI. Many disciplines (including knowledge representation, decision support, database management, prediction and planning, diagnosis and explanation, and so on) are attacking different facets of the time problem and developing new methodologies and algorithms. For example, TimeML (see <http://www.timeml.org>) is a standard mark-up scheme for temporal information. Figure 1 (on the next page) is a TimeML visualization of a sample TimeBank document—TimeBank is a directory of news articles marked-up in TimeML, see <http://timeml.org/site/timebank/timebank.html>. Another example is DAML-Time (see <http://www.cs.rochester.edu/~ferguson/daml/>),

## Time plays a key role in all aspects of the intelligence analysis process.



1 A TimeBank example, marked-up in TimeML using MITRE’s Callisto and visualized through MITRE’s Tango (see <http://timeml.org/site/tango/tool.html>).

the community to imagine better ways to visualize temporal data.

The left-most column of the diagram in Figure 2, illustrating the bank names, represents entities that have temporal information associated with them in the article. Each entity has the ability of showing or hiding its detail (controlled with the + or – symbols on the left-hand side of the entity name). Below the entity name are two intensity bars. The gray bar represents the amount of temporal uncertainty associated with the entity. The red bar represents the amount of temporal conflict within the entity. Across the top of the diagram, a timeline shows the relative occurrence of the temporal entities (a sequence is shown as black dots under the timeline). Each entity can contain a relative timeline within its space. This allows the entity to have a change in temporal zoom relative to its container timeline (see Banner Bank) or allows recording of unnormalized events within their known temporal reference frame (see Carter Bank).

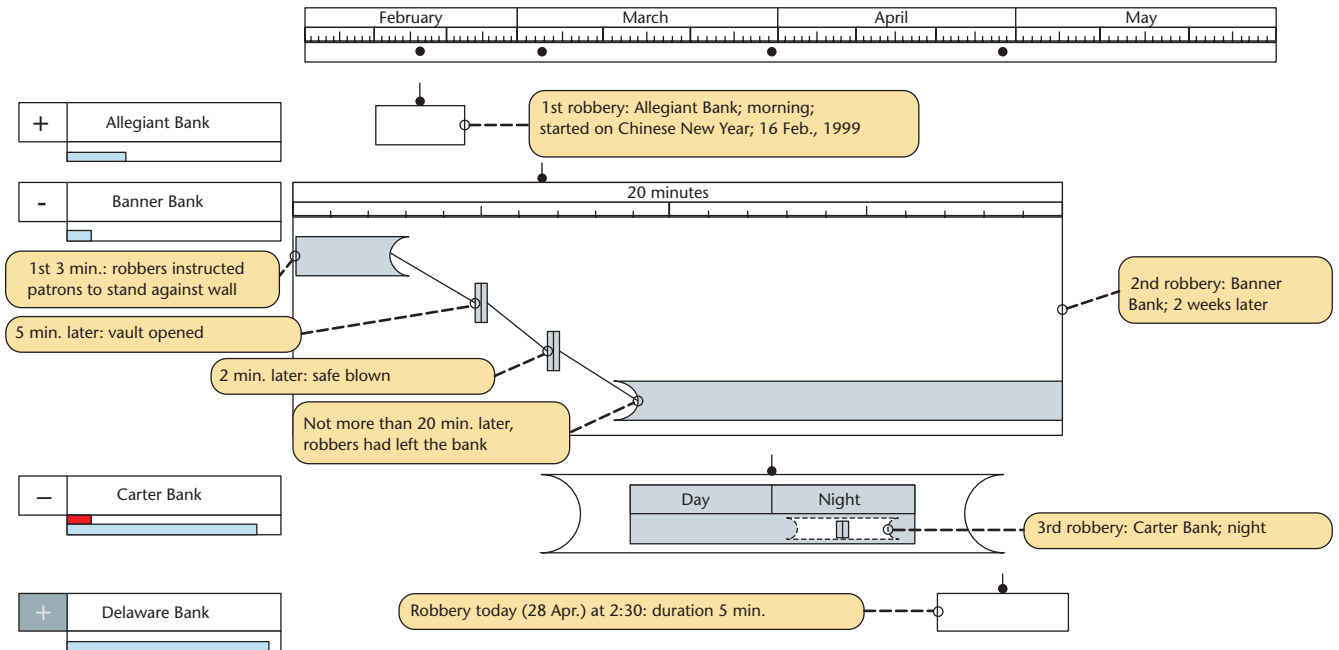
which aims to develop a representative ontology of time that covers the basic topological temporal relations on instants and intervals, measures of duration, and the clock and calendar. This enables the statement of and reasoning about temporal aspects of unstructured text documents. There is a vast amount of literature on temporal representation (such as temporal logic) and reasoning (temporal problem solvers). Since we are dealing with human language, temporal knowledge can be uncertain and incomplete, leading to issues like ambiguity, conflict, uncertainty, and so on.

Many of you will still argue that this field of extracting information is young and immature. Let us make an assumption: If the best information extractor can extract 30 percent of the information, including temporal events, then what sort of visualizations can we provide to aid users in understanding that 30 percent? What can the computer graphics community do to use that 30 percent? This is where TAD comes into the picture—this research community can provide the visualization and interaction methodologies to aid analysts in their cognitive understanding when dealing with time in their discourse. For example, Figure 2 is a more elaborate visualization of our simple bank robberies example. This visualization is a rough idea (with its corresponding interactions), but we hope it gets across the points that we are trying to make in this article about treating time as a separate and special variable. In this visualization, we try to convey the ideas of duration (instances, exact durations, and fuzzy durations), differing timescales, some information on uncertainty and conflicts, and dependencies. A timeline visualization might not always be the appropriate way to visualize data; we encourage

Temporal containers can be rectangular or concave. Rectangular containers represent known start and end times, while concavity on an edge represents uncertainty in a temporal anchor.

Consider the Carter Bank entity (starting from the outside moving inward). The first concave layer represents uncertainty of the time of the third robbery. We know it happened sometime between the second and fourth robberies. The concave portions of the outer bar extend to the second and fourth robbery markers to indicate this uncertainty. The second layer sets up a new temporal reference frame. This temporal region is divided qualitatively into day and night. By adding an event to the night interval, the diagram will convey that the event occurred at night sometime between the second and fourth robberies. The third layer of the Carter Bank temporal entity is a concave bar with a discrete time window inside. This describes that we don’t know at what time the robbery occurred, but we do know that it must have been a short time. One interesting note is being able to visualize what is missing (like not hearing the bank alarm). Missing information can provide clues in understanding a particular event and determining the next clues.

Analysts see time as an important variable that they must deal with more directly to generate a defensible knowledge product. Technically, when an analyst must defend their reasoning product, he or she is actually defending their reasoning and cognitive processes used to formulate the product. Thus, we need to identify and support different mental models, metaphors, and other devices that humans use in analyzing temporally related problems. To do this, we need to integrate the knowl-



## 2 Elaborate visualization of the simple example.

edge and methods developed by researchers in the cognitive sciences, semiotics, communications, media theory, and linguistics.

### Conclusion

The path to the truly effective use of time in analytic discourse contains many challenging obstacles, such as what metaphors can be used to describe it and what is its optimal representation in different visualizations. Today's analytic tools fall short of satisfying temporal reasoning needs of analysts. For example, there are few widely accepted methodologies appropriate for temporal dependency representation. Uncertainty and ambiguity representation standards for temporal data are almost nonexistent. The visualization community should embrace and foster standards for visual presentation for these facets of time; otherwise, exploitation of the time dimension will continue to be curtailed by the implicit representation of time found in many visualizations.

Analysts need to work with time in an unhindered way. Common analytical activities, such as evidence chaining, hypothesis testing, and scenario testing often depend heavily upon the explicit use of time (as perceived through various metaphors). If we work to further the development of temporal analytic discourse, we can expect to see immense breakthroughs in many application areas. ■


### References

1. D. Peuquet, *Representations of Time and Space*, Guildford Press, 2002.
2. B. Dowden, "Time," *The Internet Encyclopedia of Philosophy*, 2005; <http://www.utm.edu/research/iep/t/time.htm>.

3. B. Dowden, "Time Supplement," *The Internet Encyclopedia of Philosophy*, 2005; <http://www.iep.utm.edu/ancillaries/time-sup.htm#H10>.

Readers may contact Olga A. Kuchar at [Olga.Kuchar@pnl.gov](mailto:Olga.Kuchar@pnl.gov).

Readers may contact Miguel Encarnaç o at [me@imedia-labs.com](mailto:me@imedia-labs.com).




## Sign Up Today for the IEEE Computer Society's e-News

**Be alerted to**

- articles and special issues
- conference news
- registration deadlines

**Available for FREE to members.**



**computer.org/e-News**