An Analyst's Geospatial and Ontological Assistant

Albert Esterline, Roland Johnson, Edward Carr, and William Wright

Department of Computer Science, North Carolina A&T State Univ.,
Greensboro, NC, USA
esterlin@ncat.edu, roland.ncat@gmail.com,
carr@redux.comp.ncat.edu, wtwright@ncat.edu

Abstract. We discuss an Intelligence Analyst's Geospatial and Ontological Assistant (IAGOA) under development that associates an intelligence analyst's understanding of an agent's activities with the geospatial features of the area of operation where they take place. Activities are identified with frames for the corresponding verbs from the FrameNet lexical database. A modeler, using the FrameNet OWL distribution, produces software used by the analyst to update a KML file with annotations identifying instantiations of the frames elements of the relevant frames. The Google Earth API is used for rendering KML files and scripting. The agent is tracked and the analyst's conjecture of its activity is simulated; the analyst can redo her conjecture if need be. IAGOA's FrameNet-based approach instantiates concepts inherent in language, making explicit the activities and the constellation of role-fillers involved in these activities.

Keywords: geospatial ontology, FrameNet, KML, intelligence analysis

1 Introduction

We discuss an Intelligence Analyst's Geospatial and Ontological Assistant (IAGOA) under development that will associate an intelligence analyst's understanding of a player's activities with the geospatial features of the area of operation where they occur. (The neutral term "player" refers to the monitored agent, which could be a person or, e.g., an unmanned vehicle.) Fig. 1 shows a hiker ("Theme") moving along a path from his source to the goal, to which the analyst conjectures he is headed. The map already highlights geospatial features significant for human activities. They are like Gibson's affordances [1], which are clues in the environment that indicate possibilities for action, but here the clues are on the map. We have clues of what can (cf. the tracks) and what cannot be done (cf. the peaks and streams), and we see a desirable place (cf. the building). The features have different dimensionalities. We think of tracks and streams as one-dimensional since we are not interested in their widths except insofar as they impede or facilitate motion, information that could be attached as attributes. Where tracks and streams end or meet are effectively zero-dimensional landmarks as are the building and where a track terminates.

Actual activities are expressed in the first instance by verbs, but a verb by itself does not have total meaning until placed into context with its "frame elements" (or FEs), which are roles filled by the denotations of the verb's complements [2]. For example, *move* evokes the *Motion* frame, which characteristically

includes the roles *Theme* (the thing moving), *Source* (where it starts), *Path* (along which it moves), and *Goal* (where it ends) as seen in Fig. 1. (Frames in the current sense, due to Fillmore, should not be confused with those introduced by Minsky, which have become a standard AI representation.) In the current domain, the roles are often expected to be filled with geospatial features. The case shown in Fig. 1 involves understanding not only the (im)possibilities and attractions afforded by the geospatial features and what it means to move but also how the *Motion* frame is instantiated in the depicted area of operation. Here the analyst has observed the player starting his motion at the track terminus, which fills the *Source* role, and has observed him moving along the track to the location shown. The building is the obvious thing of interest, so a good guess has it fill the *Goal* role. We already know part of the path and have a good idea of how it will continue, so concatenating the segments gives a good guess for what fills the *Path* role. Some roles, however, are temporal (e.g., *Motion* includes *Speed* or *Duration*) and some are non-spatiotemporal, such as *Manner* (e.g., *stealthily*).
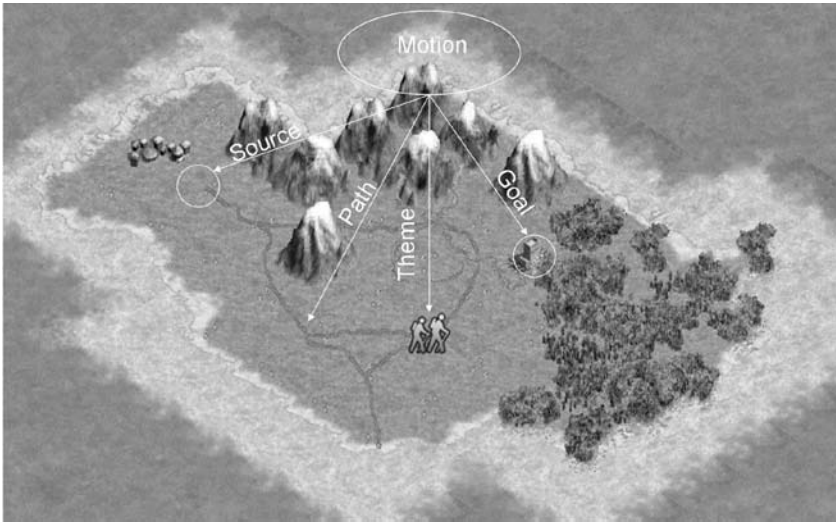

Fig. 1. A hiker ("Theme") moving along a path from his source to the goal along a path.

In IAGOA, the analyst identifies the fillers for the roles and the system simulates the player and updates the display showing what really transpires. If the conjecture deviates, the analysts can formulate a new conjecture. When the player is done, the analyst updates the case to describe what she thinks has just transpired, and this is remembered so that, when a similar case arises in the future, it may be consulted.

For visualization, we use KML, an XML-based language and Open Geospatial Consortium (OGC) standard [3] for geographic annotation and visualization by Internet-based Earth browsers. KML can represent points, sequences of connected line segments, and polygons, which can be annotated with attribute-value pairs. Location is in terms of latitude and longitude. We use the Google Earth API to render a KML file on a background of a satellite image and for scripting. The Geography Markup Language (GML) [4] is another XML-based

OGC standard for expressing geographical features. It complements KML: GML is a language to encode geographic content for any application, while KML is a visualization language.

For representing objects in the real world, we endorse the perspective of Smith's mereotopology [5], which avoids point-set topology's dependence on points that exist independently of all that is around them, and we accept Brentano's thesis that boundaries do not exist independently of the entities they bound. We thus introduce 1+Ds, which have significant length but whose width can generally be ignored, and 0+Ds, whose dimensions are so small that they are not considered in the model at hand. (In this nomenclature, a 0D is a mathematical point, a 1D a curve, and a 2D a region.) We, however, recognize the importance of fiat objects, which are established by stipulation and are not under the same mereological constraints as natural objects. Fiat 0Ds (e.g., a stipulated starting point) as well as natural 1Ds (e.g., region boundaries) and fiat 1Ds (e.g., a line of advance) are also important and, in the case of the 1Ds, help demarcate 0Ds by intersection.

We use the FrameNet online lexical database [6, 7] for frames that have been formulated by analyzing language corpora. The FrameNet project has recorded the information on frames in an OWL file [8], OWL being the Web Ontology Language [9]. An ontology is a formal, explicit specification of a shared conceptualization [10]. In IAGOA, a modeler, assisted by a program that takes as input the FrameNet OWL file, develops form elements and associated script to query the analysts for verbs and the FEs of the associated frames.

The most popular lexical database, and the most influential for ontology development, is WordNet [11]. It groups words into sets of synonyms called synsets; relations between synsets are justified psychologically by how humans process language and depend on the type of word. For example, nouns are related by subclass-of and part-of relations and their inverses. WordNet's subclass-of relations align well with the taxonomic backbone of most upper (foundation) ontologies. An upper ontology describes general concepts that are the same across all domains and supports broad semantic interoperability between ontologies that specialize it. The Suggested Upper Merged Ontology (SUMO) [12], for example, has mappings to all of WordNet.

Several ontology-based systems to assist intelligence analysts have been reported in the literature. [13] describes a Geospatial and Visual Information Ontology for analyst-specific information processing that uses WordNet as its semantic base and defines a set of properties for each lexical category. [14] describes an imagery analysis environment where observations are recorded as structured annotations using an ontology based on OWL-Time, GeoRSS, BFO, and the Defense Geospatial Information Working Group Feature and Attribute Data Registry. See [15] for a survey of ontologies relevant to geospatial intelligence. An advantage of IAGOA's FrameNet-based approach is that it instantiates concepts inherent in language, making explicit the activities and the constellation of role-fillers involved in these activities. To capture such structure, WordNet and the common upper ontologies require additional structure (such as Minsky-type frames) beyond the classes themselves.

The next section describes how we use FrameNet as a basis for an ontology of activities. The following section outlines the IAGOA system, and the final one concludes. Space restrictions prevent a complete description of IAGOA; in particular, how previous cases similar to the current case are found is not covered.

2 Ontologies for Activities

Given a word and universal grammar (UG) [16], the meaning of the word cannot be deduced without access to all essential knowledge that relates to that word. Theta theory (Θ-theory) mandates that verbs are referred to as predicates to include not only the verb but also its semantic categories, that is, thematic roles (Θ-roles). For example, a verb such as *hug* brings in two Θ-roles, an agent (subject participant) and patient (object participant).

To relate linguistic semantics to encyclopedic knowledge and utilize the concepts similar to those presented in Θ-theory, we use FrameNet [6, 7], accessible online at http://framenet.icsi.berkeley.edu/ although the data for this research is obtained mostly from FrameNet's OWL distribution [8]. FrameNet is based on Fillmore's frame semantics [17], where a frame "is any system of concepts related in such a way that to understand any one concept it is necessary to understand the entire system …" [18]. These concepts are FrameNet's frame elements (FEs), the equivalents of Θ-roles. FEs have inheritance relations, which eventually lead to semantic types, which themselves have inheritance relations. The current FrameNet OWL distribution has a meager set of semantic types. The FrameNet team suggests something like SUMO classes could be used for semantic types, and, for nouns denoting artifacts and natural kinds, they largely defer to WordNet (with its extensive hierarchical coverage of such areas) (see [7], sec. 1.1). Although a given frame may be associated with several parts of speech (e.g., the verb *move* and the noun *motion* elicit the same frame), verbs are paramount in frame semantics, and we largely restrict our attention to them. The lexical database essentially consists of lexical units (LUs, words) and associated frames. One word may be paired with several frames (polysemy), and one frame is typically associated with several words. The verb by itself does not have total meaning until it is placed into context with its frame elements. A word evokes a frame of semantic knowledge relating to the specific concept it highlights.

The instantiation of an FE might in turn be a frame with its own FEs. For example, frame *Verification* has two core FEs, *Inspector* and *Unconfirmed_content*, which is an open proposition that the *Inspector* decides by examining evidence. The *Unconfirmed_content*, being a proposition, invokes a frame. For example, consider "The player makes sure that the bridge is on the road." This evokes the *Verification* frame, with *Inspector* instantiated with the player and *Unconfirmed_content* instantiated with the proposition expressed by "The truck is on the road." The latter proposition evokes the *Being_located* frame, whose core FEs are *Theme* (instantiated with the bridge) and *Location* (expressed by "on the road"). Although the embedding of frames is most obvious in sentences with subordinate clauses, the phenomenon does not depend on that kind of syntax.

3 Software for Analysts and Modelers

IAGOA has a component for the analyst and a component for a modeler who is responsible for HTML and associated JavaScript that allow the analyst to capture her current conjecture of the player's behavior in terms of verbs and the FEs of the frames evoked by these verbs. The analyst works on a client. The

server provides HTML documents that fetch updated KML documents. The server updates a KML document according to form data sent from the client. It also constructs HTML documents that provide form data according to the verbs and FEs currently selected. The analyst can track the player and, at the same time, watch her conjecture.

IAGOA begins with a set of KML documents representing possible areas of operation that have been prepared to facilitate the sort of analysis intended. The features of interest are encoded as KML `Point` elements (0+Ds), `LineString` elements (1+Ds), and Polygon elements (2Ds). Fiat 0Ds and 1Ds are handled similarly. (KML allows only straight line segments, and the only 2D geometries are polygons; curvilinear features must be approximated.) Typed name-value pairs are added to the content of these elements using KML `ExtendData` elements. These appear in the balloon when the corresponding icon is clicked and include such attributes as the height and other salient properties of a building (a 1+D) and the width and surface type of a road (a 1+D). KML styles are used so that the designated features are readily apparent; different styles reflect different values for key attributes.

We present the design for the analyst's software before that for the modeler's software since the modeler's software makes sense only in that context.

3.1 The Analyst

The analyst is presented with a display produced on a webpage using the Google Earth API. It is assumed that data are provided (say, from a satellite) that update in near real-time the location of a player of interest[1]. When the analyst clicks near the player, she will be presented with a list of verbs. She selects one verb at a time; additional verbs can be selected after information for the selected verb is submitted and the server returns a new HTML document that fetches an updated KML document (as mentioned below). After selecting a verb, she will be presented with the associated form, which will be filled in principally with values of a spatial and sometimes temporal nature. The Google Earth API is used to identify spatial features. With a given form element in focus, the analyst clicks the feature on the KML rendering to get its id. If she wishes to select a 1+D not already defined in the KML document, she can define it by clicking a sequence of 0+Ds. Similarly, she can define a missing 2D with 1+Ds (or 1Ds) and, if necessary, 0+Ds (or 0Ds). If need be, she can define landmarks (0+Ds) on the fly. Some FEs come in alternative sets. For example, *Motion* is described in terms of not only the *Source* but also either the *Path* and *Goal* or the *Direction* and *Distance*. Since we are talking about activities, the relevant notion of time is that of *Duration*. In the case of *Motion*, if *Distance* is indicated, then one can deduce the *Duration* from the *Speed*. For aspects that are neither spatial nor temporal, the analyst sees a menu of ontological categories.

When the form is submitted, the server uses Jena (a framework for building Semantic Web applications—see http://openjena.org/) to update a copy of the KML file. It identifies the instantiations of the FEs by associating with them `ExtendData` elements that identify what FEs they instantiate. Any new

---

[1] Google Earth allows track data to be imported from some GPS devices, and some companies (e.g., Safe Software) provide spatial ETL (extract, transform and load) technology for interoperability of spatial data.

landmarks are added to the KML file. The file returned is an HTML document that fetches the new KML file. The player is again tracked on the rendering of this document, but now the conjectured change over time (such as change of position for *Move* or change in direction for *Survey*) is also shown. Also, the analyst may select another verb from the menu. A special case is where one or more FEs are instantiated with frames. The analyst selects the value for such an FE from the menu of verbs, thereby selecting the associated frame. The server's responses then present in succession the forms for the FE-instantiating frames without the analyst selecting the corresponding verbs.

The new HTML document has buttons in case the analyst decides to modify her conjecture. There are buttons for each verb and each FE associated with each verb's frame, allowing the analyst to remove an FE or entire frame, replace it, or modify it. There is also a button that returns the original KML document, throwing away all the FrameNet-related information that has been added.

The progress of certain activities can be shown by moving an indicator on the KML rendering of the area of operation. For example, motion can be shown as the player changing positions over time, and looking at (frame *Perception_active*) can be shown as the direction changing over time. This can be captured in KML using `TimeStamp` elements with the elements representing the entity at different locations at different times. Generally, only locations corresponding to marked landmarks will be used, although it might be necessary to define new landmarks (such as bends in a road). This time-varying information can be captured in several ways.

When a case that the analyst has been following is resolved, she saves this case so that it may be consulted for suggestions when a similar case arises. IAGOA allows the analyst to record some of her understanding of the case. When she saves the final state, she selects those additional features deemed significant, and they are annotated with `ExtendData` elements. The analyst can also indicate which differences in length are significant and which are not.

3.2 The Modeler

The modeler, using FrameNet resources and several programs, constructs, for each verb the analyst may choose, the HTML document (and accompanying script) whose form is presented by the server to the analyst. The same HTML document is used with any number of KML documents representing areas of operation.

XSLT is used to extract the lexical units and associated frames from the FrameNet HTML documents that contain these associations. This produces an HTML document. When it is rendered, the modeler selects the lexical unit he wishes to capture. If there is more than one frame associated with that lexical unit, he must select among them; he can view the FrameNet HTML pages to see their definitions. Selecting a frame will invoke a Jena program that processes the FrameNet OWL file, starting with that frame. The modeler will be presented with one FE after another associated with that frame. He will select the FEs of interest (again referring to the HTML pages for explanations). For each selected FE, the Jena program will chase back the inheritance relations until it gets to the semantic types, and the modeler can follow the inheritance relations among semantic types.

For a given frame, what FEs are provided and which are required and in what combination are questions that must be addressed and can result in considerable structure in the choices presented to the analyst. According to sec. 3.2.1 of [7], a core FE normally is instantiated with its frame, but it can often be omitted. As noted, some FEs form alternative sets. Also, a core FE may be "null instantiated" (see [7], sec. 3.2.3), as when the element is understood in the context; note that the FE instantiated with the player can generally be null instantiated. There are also several relations among FEs (see [7], sec. 3.2.2). Sometimes the occurrence of one core FE requires the occurrence of another, while some FEs are incompatible.

Java code executed with the Jena program will construct the HTML form from the selected FEs with considerable help from the analyst. The modeler will control any explanatory text. If the semantic type of an FE is spatial, then the form will be equipped with the code required to capture mouse clicks on the Google Earth display. If the semantic type of an FE is a duration or speed, then the form will be equipped with appropriate text boxes and menus for the units, and appropriate validation code will be included. If the semantic type of an FE is non-spatiotemporal, then the modeler will populate a menu with values consistent with the semantic type. The program must also produce HTML fragments and corresponding JavaScript code that the server will use in constructing HTML documents that form the response when the analyst updates the current case. These are simply fragments of the HTML document just discussed.

4 Conclusion

We have discussed the Intelligence Analyst's Geospatial and Ontological Assistant (IAGOA), which is under development and will associate an intelligence analyst's understanding of a player's activities with the geospatial features of the area of operation where they take place. For visualization, the Google Earth API is used to render a KML file on a background of a satellite image and for scripting. IAGOA has a component for the analyst and a component for a modeler who is responsible for code that allows the analyst to capture her current conjecture of the player's behavior in terms of verbs and the FEs (frame elements) associated with the FrameNet frames evoked by these verbs. The principal input to the modeler software is the FrameNet OWL file. The analyst identifies the fillers for the FEs and IAGOA simulates the player and also updates the display showing what really transpires. If the conjecture turns out false, the analyst tries another. When the player is done, the analyst updates the case, which is remembered for future comparisons.

The aspects of IAGOA not covered in this paper mostly relate to the use of qualitative representations of spatial properties and relations. (See our presentation in [19].) IAGOA uses qualitative techniques to find stored cases similar to the current case and to gain insight into the player's perspective and opportunities. FrameNet provides suggestions for numerous enhancements to IAGOA. For example, it uses several frame-to-frame relations to situate frames in semantic space (see [7], chap. 6), and IAGOA could follow these relations to extend its dialog with the analyst. Among the aspects that are beyond the scope of a system like IAGOA without major enhancement, the most notable is the handling of multiple players.

References

1. Gibson, J.J.: The Ecological Approach To Visual Perception. Psychology Press, London (1986)
2. Esterline, A. and BouSaba, C.: An Ontology for Tactical Behaviors Derived from Verb Frames. In: Proc. of the Unmanned Systems Technology X Conf. (SPIE Defense and Security Symp.), Orlando, FL, March (2008)
3. OGC, KML, http://www.opengeospatial.org/standards/kml
4. OGC, Geography Markup Language (2010) http://www.opengeospatial.org/standards/gml
5. Smith, B.: Mereotopology: A Theory of Parts and Boundaries. Data and Knowledge Engineering, vol. 20, pp. 287--303 (1996)
6. Boas, H.: From theory to practice: frame semantics and the design of FrameNet. In: Langer, S. and Schnorbusch, D. (eds.) Semantisches Wissen im Lexikon.: Gunter Narr Verlag, Tübingen, Germany. pp. 129--160 (2005)
7. Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., and Scheffczyk, J. FrameNet II: Extended Theory and Practice, Web Publication (2006) http://framenet.icsi.berkeley.edu/book/book.html.
8. J. Scheffczyk, J., Baker, C.F., and Narayanan, S.: Ontology-Based Reasoning about Lexical Resources. In: Proceedings of the Workshop on Interfacing Ontologies and Lexical Resources for Semantic Web Technologies (OntoLex 2006), Genoa, Italy, pp. 1--8 (2006)
9. Smith, M., Welty, C., and McGuinness, D.: OWL Web Ontology Language Guide, W3C Recommendation. (2004) http://www.w3.org/TR/owl-guide/.
10. Guarino, N.: Formal ontology and information systems. In: Guarino, N. (ed.). Formal Ontology and Information Systems, IOS Press. (1998)
11. Fellbaum, C.: (ed.): WordNet: An Electronic Lexical Database, The MIT Press. (1998)
12. Niles, I. and Pease, A.: Towards a Standard Upper Ontology. In: Welty, C. and Smith, B. (eds) Proc. 2nd Int. Conf. on Formal Ontology in Information Systems (FOIS-2001), Ogunquit, MA (2001).
13. Basu, C., Cheng, H., and Fellbaum, C.: Creating a Geospatial and Visual Information Ontology for Analysts. In: K. S. Hornsby (ed.) Proc. 2nd Int. Ontology for the Intelligence Community (OIC-2007), Columbia, MD (2007)
14. Self, T., Kolas, D., Dean, M.: Ontology-Driven Imagery Analysis. In: K. S. Hornsby (ed.) Proc. 2nd Int. Ontology for the Intelligence Community (OIC-2007), Columbia, MD (2007)
15. Ressler, J. and Dean, M.: Geospatial Ontology Trade Study. In: K. S. Hornsby (ed.), Proc. 2nd Int. Ontology for the Intelligence Community (OIC-2007) Columbia, MD (2007)
16. Ouhalla, J.: Introducing Transformational Grammar: From Principles and Parameters to Minimalism, Hodder Arnold, London (1999)
17. Fillmore, C.J.: The case for case. In: Bach and Harms (ed.): Universals in Linguistic Theory. Holt, Rinehart, and Winston, New York, pp.1--88 (1968)
18. Petruck, M.: Frame semantics. In: Verschueren, J., Östman, J-O., Blommaert, J., and Bulcaen, C. (eds.) Handbook of Pragmatics. John Benjamins, Philadelphia (1996)
19. BouSaba, C., Esterline, A., et al.: Spatial Ontologies for Tactical Behaviors. In: Proc. Unmanned Systems Technology X Conf. (SPIE Defense and Security Symp.), Orlando, FL (2008)