

TESLA: A Tool for Annotating Geospatial Language Corpora

Nate Blaylock and Bradley Swain and James Allen

Institute for Human and Machine Cognition (IHMC)

Pensacola, Florida, USA

{blaylock, bswain, jallen}@ihmc.us

Abstract

In this paper, we present The gEoSpatial Language Annotator (TESLA)—a tool which supports human annotation of geospatial language corpora. TESLA interfaces with a GIS database for annotating grounded geospatial entities and uses Google Earth for visualization of both entity search results and evolving object and speaker position from GPS tracks. We also discuss a current annotation effort using TESLA to annotate location descriptions in a geospatial language corpus.

1 Introduction

We are interested in *geospatial language understanding*—the understanding of natural language (NL) descriptions of spatial locations, orientation, movement and paths that are grounded in the real world. Such algorithms would enable a number of applications, including automated geotagging of text and speech, robots that can follow human route instructions, and NL-description based localization.

To aide development of training and testing corpora for this area, we have built The gEoSpatial Language Annotator (TESLA)—a tool which supports the visualization and hand-annotation of both text and speech-based geospatial language corpora. TESLA can be used to create a gold-standard for training and testing geospatial language understanding algorithms by allowing the user to annotate geospatial references with object (e.g., streets, businesses, and parks) and latitude and longitude (lat/lon) coordinates. An integrated search capability to a GIS database with results presented in Google Earth allow the human annotator to easily annotate geospatial references with ground truth.

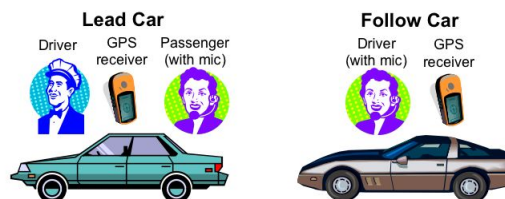


Figure 1: A session in the PURSUIT Corpus

Furthermore, TESLA supports the playback of GPS tracks of multiple objects for corpora associated with synchronized speaker or object movement, allowing the annotator to take this positional context into account. TESLA is currently being used to annotate a corpus of first-person, spoken path descriptions of car routes.

In this paper, we first briefly describe the corpus that we are annotating, which provides a grounded example of using TESLA. We then discuss the TESLA annotation tool and its use in annotating that corpus. Finally, we describe related work and our plans for future work.

2 The PURSUIT Corpus

The PURSUIT Corpus (Blaylock and Allen, 2008) is a collection of speech data in which subjects describe their path in real time (i.e., while they are traveling it) and a GPS receiver simultaneously records the actual paths taken. (These GPS tracks of the actual path can aide the annotator in determining what geospatial entities and events were meant by the speaker’s description.)

Figure 1 shows an example of the experimental setup for the corpus collection. Each session consisted of a lead car and a follow car. The driver of the

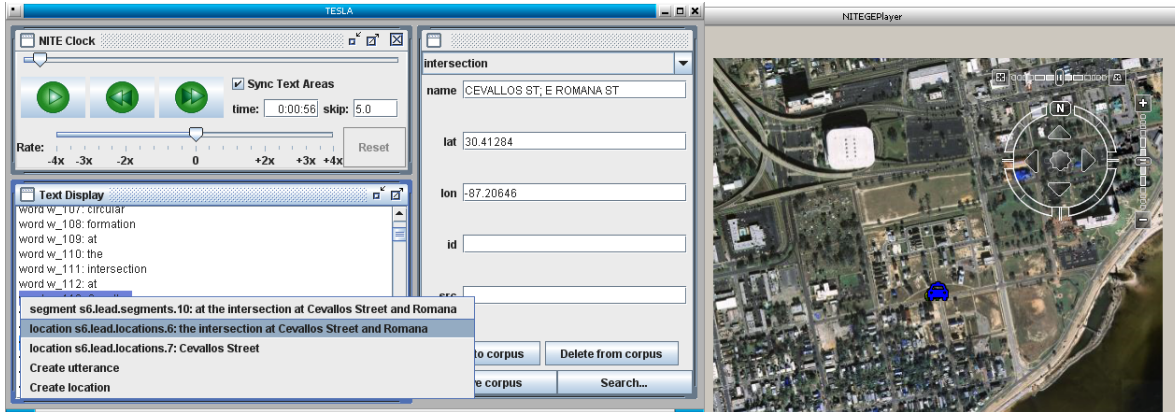


Figure 2: The TESLA annotation and visualization windows

lead car was instructed to drive wherever he wanted for an approximate amount of time (around 15 minutes). The driver of the follow car was instructed to follow the lead car. One person in the lead car (usually a passenger) and one person in the follow car (usually the driver) were given close-speaking headset microphones and instructed to describe, during the ride, where the lead car was going, as if they were speaking to someone in a remote location who was trying to follow the car on a map. The speakers were also instructed to try to be verbose, and that they did not need to restrict themselves to street names—they could use businesses, landmarks, or whatever was natural. Both speakers’ speech was recorded during the session. In addition, a GPS receiver was placed in each car and the GPS track was recorded at a high sampling rate. The corpus consists of 13 audio recordings¹ of seven paths along with the corresponding GPS tracks. The average session length was 19 minutes.

3 TESLA

TESLA is an extensible tool for geospatial language annotation and visualization. It is built on the NXT Toolkit (Carletta et al., 2003) and data model (Carletta et al., 2005) and uses Google Earth for visualization. It supports geospatial entity search using the TerraFly GIS database (Rishe et al., 2005). Currently, TESLA supports annotation of geospatial location referring expressions, but is designed to be easily extended to other annotation tasks for geospa-

¹In one session, there was no speaker in the lead car.

tial language corpora. (Our plans for extensions are described in Section 6.)

Figure 2 shows a screenshot of the main view in the TESLA annotator, showing a session of the PURSUIT Corpus. In the top-left corner is a widget with playback controls for the session. This provides synchronized playback of the speech and GPS tracks. When the session is playing, audio from a single speaker (lead or follow) is played back, and the blue car icon in the Google Earth window on the right moves in synchronized fashion. Although this Google Earth playback is somewhat analogous to a video of the movement, Google Earth remains usable and the user can move the display or zoom in and out as desired. If location annotations have previously been made, these pop up at the given lat/lon as they are mentioned in the audio, allowing the annotator to verify that the location has been correctly annotated. In the center, on the left-hand side is a display of the audio transcription, which also moves in sync with the audio and Google Earth visualization. The user creates an annotation by highlighting a group of words, and choosing the appropriate type of annotation. The currently selected annotation appears to the right where the corresponding geospatial entity information (e.g., name, address, lat/lon) can be entered by hand, or by searching for the entity in a GIS database.

3.1 GIS Search and Visualization

In addition to allowing information on annotated geospatial entities to be entered by hand, TESLA also supports search with a GIS database. Cur-

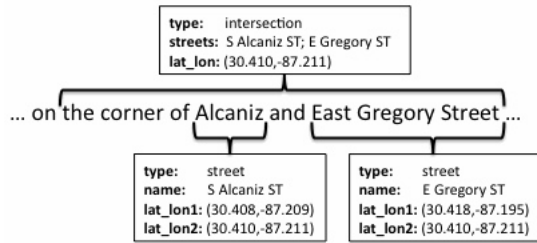


Figure 4: Sample annotations of referring expressions to geospatial locations

streets and an intersection. Here the intersection referring expression spans two referring expressions to streets, and each is annotated with a canonical name as well as lat/lon coordinates. Note also that our annotation schema allows us to annotate embedded references (here the streets within the intersection).

5 Related Work

The SpatialML module for the Callisto annotator (Mani et al., 2008) was designed for human annotation of geospatial locations with ground truth by looking up targets in a gazetteer. It does not, however, have a geographic visualization components such as Google Earth and does not support GPS track playback.

The TAME annotator (Leidner, 2004) is a similar tool, supporting hand annotation of toponym references by gazetteer lookup. It too does not, as far as we are aware, have a visualization component nor GPS track information, likely because the level of geospatial entities being looked at were at the city/state/country level. The PURSUIT Corpus mostly contains references to geospatial entities at a sub-city level, which may introduce more uncertainty as to the intended referent.

6 Conclusion and Future Work

In this paper, we have presented TESLA—a general human annotation tool for geospatial language. TESLA uses a GIS database, GPS tracks, and Google Earth to allow a user to annotate references to geospatial entities. We also discussed how TESLA is being used to annotate a corpus of spoken path descriptions.

Though currently we are only annotating PURSUIT with location references, future plans in-

clude extending TESLA to support the annotation of movement, orientation, and path descriptions. We also plan to use this corpus as test and training data for algorithms to automatically annotate such information.

Finally, the path descriptions in the PURSUIT Corpus were all done from a first-person, ground-level perspective. As TESLA allows us to replay the actual routes from GPS tracks within Google Earth, we believe we could use this tool to gather more spoken descriptions of the paths from an aerial perspective from different subjects. This would give us several more versions of descriptions of the same path and allow the comparison of descriptions from the two different perspectives.

References

- Nate Blaylock and James Allen. 2008. Real-time path descriptions grounded with gps tracks: a preliminary report. In *LREC Workshop on Methodologies and Resources for Processing Spatial Language*, pages 25–27, Marrakech, Morocco, May 31.
- Jean Carletta, Stefan Evert, Ulrich Heid, Jonathan Kilgour, Judy Robertson, and Holger Voormann. 2003. The NITE XML toolkit: flexible annotation for multimodal language data. *Behavior Research Methods, Instruments, and Computers*, 35(3):353–363.
- Jean Carletta, Stefan Evert, Ulrich Heid, and Jonathan Kilgour. 2005. The NITE XML toolkit: data model and query language. *Language Resources and Evaluation Journal*, 39(4):313–334.
- Jochen L. Leidner. 2004. Towards a reference corpus for automatic toponym resolution evaluation. In *Workshop on Geographic Information Retrieval*, Sheffield, UK.
- Inderjeet Mani, Janet Hitzeman, Justin Richer, Dave Harris, Rob Quimby, and Ben Wellner. 2008. SpatialML: Annotation scheme, corpora, and tools. In *6th International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Morocco, May.
- N. Rishe, M. Gutierrez, A. Selivonenko, and S. Graham. 2005. TerraFly: A tool for visualizing and dispensing geospatial data. *Imaging Notes*, 20(2):22–23.