BUILDING ROUTE BASED MAPS FOR THE VISUALLY IMPAIRED FROM NATURAL LANGUAGE ROUTE DESCRIPTIONS

John Nicholson
Department of Computer Science
Old Main 405
Utah State University
Logan, UT, USA, 84322-4205
john.nicholson@aggiemail.usu.edu

Vladimir Kulyukin
Department of Computer Science
Old Main 401E
Utah State University
Logan, UT, USA, 84322-4205
Vladimir.Kulyukin@usu.edu

Jim Marston
Department of Geography
1832 Ellison Hall
UC Santa Barbara
Santa Barbara, CA, USA 93106-4060
marstonj@geog.ucsb.edu

Abstract

Introduction: Maps are a common method for providing spatial information and learning about indoor and outdoor areas. Street maps and building maps enable people to find routes to previously unvisited locations. Unfortunately, because these maps are primarily visual in nature, people with visual impairments often cannot take advantage of the information available on these maps. Two mapping options available to people with visual impairments are tactile maps and GPS-based maps, but even these maps cannot always provide blind travelers with operational route knowledge.

Objectives: One source of spatial information, which can augment the current mapping solutions, is natural language route descriptions provided by independent blind travelers who are familiar with the target area. For a traveler unfamiliar with a large area such as a university campus, a natural language route description that points out specific landmarks and spatial entities as well as provides context sensitive traveling instructions may improve the traveler’s use of a tactile or GPS-based map. Because natural language is unstructured, our long-term goal is to develop a system for building a route-based topological map from a collection of free text route descriptions by independent blind travelers. This paper focuses on finding and extracting landmarks from free text route descriptions written in English. The paper provides a brief explanation of how a topological map built from extracted landmarks can be used to generate new natural language route descriptions.

Methodology: We developed a web-based questionnaire that asked people with visual impairments to write descriptions of routes with which they were familiar. Each respondent provided an indoor route and an outdoor route. We received 52 responses, resulting in 104 English route descriptions. To evaluate the landmark extraction process, each set of route descriptions was divided into two groups. 18 randomly selected route descriptions from each set, indoor and outdoor, were assigned to an evaluation group and the remaining route descriptions were assigned to a training group, resulting in 36 descriptions in the evaluation group and 68 descriptions in the training group. The training set of route descriptions was analyzed for linguistic patterns that are used to identify landmarks. From these patterns, rules were written in the Java Annotations Pattern Engine (JAPE) pattern language. When a rule finds text that matches its pattern, it extracts the portion of the text containing the landmark.

Results: The landmarks in the evaluation set of route descriptions were first identified manually. The JAPE pattern rules were run against the evaluation set. The landmarks extracted by the JAPE rules were then compared against the manually identified routes, and the standard information extraction scores (recall, precision, and F-measure) were computed. For the combined set of indoor and outdoor route descriptions the recall for the extraction process was 0.8487, the precision was 0.8286, and the F-measure was 0.8386.

Conclusions: Although the landmark extraction process did not extract several landmarks and incorrectly identified some text segments as landmarks, the system correctly extracted the majority of the manually identified landmarks. The extraction errors were due to the fact that the route descriptions contained incorrect spelling and grammar and had a wide variety of writing styles. It is unlikely that all possible linguistic patterns that identify landmarks would be seen in only 68 training descriptions. However, the results demonstrate that landmarks can be automatically extracted from natural language route descriptions.

Introduction

Because traditional street and building maps are primarily visual in nature, people with visual impairments often cannot take advantage of the maps' information available on these maps. Research continues into developing mapping solutions for the visually impaired with solutions including tactile maps and GPS-annotated maps. Both of these approaches rely on data from street level maps and often miss other information sources useful to blind travelers. One such source often missed in tactile and GPS-based maps is route descriptions by independent blind travelers who are familiar with the target area. For example, over the years, visually impaired seniors at a university will have learned of a large set of routes between buildings and room-to-room routes inside the buildings on the campus.

This knowledge can be communicated to new students through natural language route descriptions. There exists research evidence that route knowledge sharing is routinely done by people with visual impairments in that they share route descriptions and verbally guide each other over cell phones (Gaunet & Briffault 2005; Kulyukin *et al.*, 2008). Capturing the route-based knowledge of experienced travelers would allow many spatial databases to be retrofitted for travelers with visual impairments. Passini and Proulx (1988) showed that people who have visual impairments prepare more for travel, make more decisions, and use more information than sighted travelers, so a route-based map built from experienced travelers' route descriptions could present spatial information to other travelers at a level of detail that they need for successful and safe navigation. Capturing this information from free text route descriptions would enable non-technical users to share their knowledge in a format with which they are already familiar. Natural language is a convenient format for learning new routes as well.

There are several issues with free text route instructions as a spatial information source. First, natural language is unstructured information whereas path-planning algorithms need structured representations. Second, since free text route description would, most likely, be typed, grammar and spelling errors are a norm. Lastly, landmarks in natural language descriptions can be referred to in multiple ways. For example, a door could be referred to as "the door to the room," "Room 225's door", or just "the door."

To address these issues, we developed the Route Analysis Engine (RAE), a software program for extracting landmarks from natural language route descriptions and compiling the extract landmarks into sequential path data structures. New route descriptions can be generated from the set of available paths. Our paper focuses on landmark extraction and sequential path compilation. Only a brief explanation of new route description generation is given.

Landmark Autotagging and Path Inference

RAE's process of extracting landmarks from natural language route descriptions is called autotagging. Autotagging breaks a route description into a list of sentences, each of which is processed for landmarks. A sentence is tagged with all landmarks found in it. Since some sentences may not mention a landmark, a sentence can have zero or more landmark tags. Autotagging relies on information extraction (IE) techniques (Cunnigham 2006). Instead of a deep level of language understanding, we identify linguistic and grammatical patterns that people use when referencing landmarks in route descriptions. RAE's autotagging module is implemented on top of the ANNIE component of the GATE system (Cunnigham 2002). ANNIE rules are written in the regular expression based language JAPE (Java Annotations Pattern Engine). The rules define text patterns that signal the presence of a landmark.

Some landmarks are well known. For example, the buildings on a university campus have well-known names and the rooms in many buildings have unique names, such as Room 405. In order to represent these well-known landmarks, RAE refers to a hierarchical set of known landmarks. The hierarchy is a tree representing "part-of" relationships. Larger and more general landmarks are stored higher up in the hierarchy, and landmarks that are more specific are stored lower in the hierarchy.

As one moves down the hierarchy, the landmarks move from large areas, such as the entire university, down to landmarks marking specific locations, e.g. a water fountain on a specific floor of a specific building. The concept of hierarchical landmarks is similar to regions in the topological level of the spatial semantic hierarchy (SSH) (Kuipers 2000). Regions in the SSH are areas that can contain smaller regions and can be part of larger regions; landmarks in the hierarchy can have sets of landmarks as children and can be children of larger landmarks. Currently, this landmark hierarchy is managed manually. In the ideal case, when autotagging identifies a landmark, the landmark matches to a landmark in the hierarchy. When such a match cannot be found, a user can manually tag a sentence at a later time and add the newly found landmark at a proper level in the hierarchy. To demonstrate the process of autotagging, consider this fragment of an actual route description submitted by a visually impaired traveler: “You are standing with your back to the south entrance to the Quick Stop. Turn left so you are walking east. On your left you will pass the ATM machines which make distinctive sounds, and the campus post office and mailboxes...” The sentences are autotagged as follows:

1. SENTENCE: You are standing with your back to the south entrance to the Quick Stop.
   LANDMARK TAGS: <south entrance>, <Quick Stop>
2. SENTENCE: Turn left so you are walking east.
   LANDMARK TAGS: EMPTY
3. SENTENCE: On your left you will pass the ATM machines which make distinctive sounds, and the campus post office and mailboxes.
   LANDMARK TAGS: <ATM machines>, <campus post office>, <mailboxes>

After autotagging, the route description is given a unique id. Each sentence is assigned a unique, ascending id and is assigned to the route description. Each sentences’ tags are saved and a reference links the tag to the appropriate sentence. If the tag represents a landmark in the hierarchy, that relationship is maintained. A user can approve the landmark tags, delete them, modify them, or add additional tags. To rebuild the original route description, the sentences are concatenated via their unique ids in ascending order.

Autotagging Evaluation

In 2007, we placed a questionnaire on the Internet asking visually impaired travelers to submit descriptions of two routes with which they were familiar and traveled regularly (Kulyukin *et al.*, 2008). One route was an indoor route from one location in a building to another location in the same building; the other route was an outdoor route between two outdoor locations. We received 52 responses for a total of 104 route descriptions. To evaluate the landmark extraction process, each set of route descriptions were divided into two groups. 18 randomly selected route descriptions from each set, indoor and outdoor, were assigned to an evaluation group and the remaining route descriptions were assigned to a training group, resulting in 36 descriptions in the evaluation group and 68 descriptions in the training group. The route descriptions in the training group were analyzed manually for common text patterns for identifying landmarks. In analyzing route descriptions, we used the list of spatial prepositions by Jackendoff and Landau (1992). The found patterns were coded as JAPE pattern rules.

The rules are executed in four phases. In the first phrase, eight groups of keywords and prepositions that are used before or after landmarks in route directions are marked.

1. Cardinal directions, e.g., North, South, western, etc.
2. Distance related terms including terms such as “feet”, “yard” and “steps” as well as more general terms such as “length” and “width”.
3. Jackendoff's 49 simple transitive prepositions, e.g., “about”, “across”, and “in”.
4. Jackendoff's 24 intransitive prepositions, e.g., “away”, and “together”
5. Jackendoff's 8 compound transitive prepositions, e.g., “to the right of.”
6. Terms related to angles such as “angle” and “degrees”.
7. Terms and phrases representing Talmy's (1983) concept of a biased part, i.e., references to parts of an object. In the sentence, “Enter the door at the front of the building,” “building” is the object of interest and “front” is the biased part. Other sample phrases would include “at the end of” and “in the middle of”.
8. Egocentric references, i.e., references to the traveler including the traveler's cane and guide dog, such as “to your right”, “with your back”, and “your dog”.

The second phase looks for matches in the landmark hierarchy. If a text segment matches a landmark in the hierarchy, it is tagged as a landmark. A reference is retained to the landmark in the hierarchy, since this tag refers to a specific authentic landmark. The third phase marks all noun phrases in text as potential landmarks. Since all landmarks are represented as noun phrases but not all noun phrases are landmarks, the fourth and final phase heuristically annotates only the appropriate noun phrases as landmarks. Specifically, the fourth phase looks for text phrases that match patterns consisting of the markers found in phase one, the noun phrases found in phase three, and other words that join the markers and noun phrases.

Each of the eight groups in the first phase has a corresponding set of rules in the last phase. For example, the phase 1 distance group has a set of rules related to distance in phase 4. There are two exceptions to this one-to-one relationship. The first exception is the rules that use verbs as the key for identifying landmarks. Since the route instructions are primarily command and action oriented, verbs are a good identifier that some object is being referred to, as in “Follow the hall” or “Open the door”. This rule set relies on the fact that prior to running these phases, ANNIE processes the entire text with a part-of-speech tagger which annotates each word with its part-of-speech. The second exception is the rules that look for patterns similar to Talmy's (1983) concept of secondary reference objects. For example, in the sentence “Go through the entrance of the building,” the “entrance” is the primary reference object and would be considered a landmark due to the spatial intransitive “though”. However, “building” is also a landmark even though it has no spatial marker. However, its connection to “door” via the “of” marker reveals it to be a secondary landmark. This rule set requires that the other nine sets of rules in this phase be run first. As a brief example of how the landmark extraction and pattern rules work, the spatial intransitive rule set in the fourth phase includes this pattern:

```
        {SpatialIntransitive}
        ( {Token.category==IN}| {Token.category==TO} )*
        ( {Token.category==DT} )?
        ( {NounPhrase} ):landmark
```

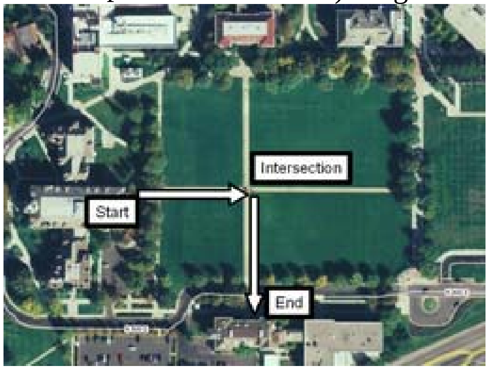
This pattern matches any text fragment in a sentence which starts with a spatial intransitive marked during the first phase, followed by 0 or more prepositions (IN and TO part-of-speech categories), followed by 0 or 1 determinants (DT part-of-speech category), and ending with a noun phrase identified during phase three. Any string of text matching this pattern would have the noun phrase tagged as a landmark. For example, in “You are in the main hall,” “in” is a spatial intransitive, “the” is a determinant, and “main hall” is a noun phrase. Since “in the main hall” can be reduced to the token sequence “<SpatialInstransitive> <DT> <NounPhrase>”, the phrase “main hall” would be identified as a landmark.

The task of creating a new route description out of two intersecting routes is done by joining appropriate segments of the intersecting route descriptions. Path inference requires a set of tagged route descriptions, a starting location, and an ending location as input. The tagged route descriptions are compiled into a directed graph and the system searches for paths in the graph between the starting and ending landmarks. If a path is found, the sentences along that path are used to create the new route description. Although space limits a full discussion of the path inference process, the following example describes the basic process. Suppose one user had entered the following route that describes a route from Animal Science building to the Ray B. West building on USU's campus (see Figure 1a): Exit the Animal Science building doors on the south side. Walk straight until you find the sidewalk entrance to the Quad's sidewalk. Walk south, passing the main intersection until you detect a road. Carefully cross the street. Continue to walk south until you find the doors to the Ray B. West building.

Another user has entered the following route that describes a route from Old Main to the Distance Learning Center on USU's campus (see Figure 1b): Exit Old Main walking east. You will walk through the Quad, passing the intersection. Keep walking straight until you run into grass. Turn left, walking north. Walk until you detect the bike racks on your right and then turn right. Walk east until you find the stairs leading to the entrance to the distance learning center.

Note that both of these descriptions mention an intersection. The first refers to it as the “main intersection” and the second as the “intersection.” During autotagging, these are extracted and the appropriate sentences are tagged. During user approval of the tags,

a) original text route description b) original text route description



c) inferred text route description

Figure 1. Visual representations of the route descriptions overlaid on a map of USU.

users would ensure that these tags referred to the known tag “Quad sidewalk intersection” representing this intersection in the hierarchy. Since both route descriptions have sentences that are tagged with a common, unique landmark from the landmark hierarchy set, it is now possible to generate a proposed new route description from Old Main to Ray B. West (see Figure 1c). The path inference would find the following new description: (From second) Exit Old Main walking east. You will walk through the Quad, passing the intersection. (From first) Walk south until you detect a road. Carefully cross the street. Continue to walk south until you find the doors to the Ray B. West building. The first half of the new route comes from the second description, while the second half from the first description. The key is that both original route descriptions have sentences tagged with the landmark “Quad sidewalk intersection.” Because the sentences are not edited but only joined during the construction of the new route description, a user is still required to ensure that the new route description is clear and can be safely followed.

Results and Discussion

After the pattern rules were developed from the training set, the rules were tested on the evaluation set of route directions. All landmarks in the evaluation set were identified

	Indoor Routes	**Outdoor Routes**	**Totals**
Correct	252	460	712
Partial	18	28	46
Missing	31	77	108
Incorrect	27	102	129

Table 1. Results of running the extraction rules on the evaluation route descriptions.

		Indoor Routes	**Outdoor Routes**	**Totals**
Precision (P)	$\frac{\text{correct + (partial * .5)}}{\text{correct + partial + incorrect}}$	0.8788	0.8034	0.8286
Recall (R)	$\frac{\text{correct + (partial * .5)}}{\text{correct + partial + missing}}$	0.8671	0.8389	0.8487
F-measure	$\frac{2.0 * P * R}{P + R}$	0.8729	0.8208	0.8386

Table 2. Computed scores for each evaluation set.

manually. Four scores were calculated:

- Correct – the autotagged and manually annotated landmarks match exactly.
- Partial – the autotagged landmark overlaps the manually annotated landmark. For example, in "From the information desk walk straight," "information desk" would be manually annotated but autotagging tagged "information desk walk".
- Missing – a landmark was manually annotated but not found during autotagging.
- Incorrect – autotagging incorrectly identified some text segment as a landmark.

These scores, shown in Table 1, were used to calculate precision, recall, and F-measure (Chinchor 1992) for each group and the total evaluation set (see Table 2). The counts and scores show that the majority of the manually annotated landmarks, 82.2%, were extracted correctly. The scores for the indoor routes are slightly higher than for the outdoor routes. This may be because the indoor routes tended to be shorter than the outdoor routes. The indoor routes have an average of 9.1 sentences and the outdoor routes have an average of 16.2 sentences. The average sentence length is comparable at 13.8 words per sentence for indoor routes and 14.0 words per sentence for outdoor routes. The extra sentences in outdoor routes may signify either that the distances for indoor routes were shorter than the distances for the outdoor routes or that outdoor environments are less structured and require more information than indoor routes to successfully navigate. In either case, further analysis is required to verify these conjectures. A number of landmarks were missing or not extracted, 12.9%, and a larger number of incorrect landmarks were autotagged. This is not surprising given the data set. The route descriptions are all in English but have a wide variety styles. Some sentences are simple and declarative, e.g., “Go out the door and make a right.” Others are longer with more complicated phrasing, e.g., “If you miss it there are benches on the outside which will tell you your near where you want to be listen for traffic though because there is a drive through you've passed the door if you get to that side of the building and need to backtrack.” Since people have a wide variety of writing styles, it is unlikely that all possible grammatical patterns would have been seen in 68 training examples. A larger training set would help resolve this issue.

Related Work

Smith-Kettlewell's Tactile Map Automated Production (TMAP) (Miele *et al.* 2006) is an example of the tactile map. The source data is the US Census Bureau’s Topologically Integrated Geographic Encoding and Referencing System (TIGER®) line data which contains features such as streets, roads, and rivers, and geographic boundaries. TMAP has also been extended to work the Talking Tactile Tablet (Miele *et al.* 2006) which provides a more interactive approach than a printed tactile map. The issue with TMAP is not the technology but with the data source. TIGER data is meant to support census reporting, not assistive navigation for the blind. The level of detail may not be sufficient in some areas for a person with a visual impairment to follow a route safely. Sendero GPS (Sendero, 2009) is an example of a GPS-based solution. While primarily an assistive device to be used during navigation, it also provides a preview feature called the "virtual explore" mode. In this mode, the user can preview and explore a route as if they were actually walking it. The system contains over 13,000,000 points of interests which allow the user to know which buildings and stores they will be walking by as they travel the route. As with all GPS systems, the main limitation of this type of technology is that is does not work indoors.

Conclusions

RAE's autotagging process can extract landmarks from natural language route descriptions. This changes the representation from unstructured to structured data. Once a set of structured route descriptions are available, path planning and inference are applicable. This allows parts of different route descriptions to be combined to form new route descriptions. User editing of the new route descriptions ensures that the routes are understandable and safe. RAE is intended to complement other mapping technologies for the visually impaired. A system like TMAP could supplement a tactile map with a natural language route direction explaining finer details. RAE is designed to not only guide travelers and build maps targeted towards people with visual impairments but also to help provide a better understanding of the types of route directions used and needed by travelers with visual impairments. A better understanding of route descriptions produced by the people with visual impairments will help to improve all navigation assistance tools.

Acknowledgments

This research has been supported, in part, through NSF grant IIS-0346880 awarded to the second author.

References

Chinchor, N., 1992. Muc-4 evaluation metrics. In Association for Computational Linguistics. *Proceedings of the 4th conference on message understanding*. Morristown, NJ, USA, 1992. pp. 22-29.

Cunningham H (2006), Information Extraction, Automatic. In K. Brown, ed. *Encyclopedia of Language & Linguistics, Second Edition*. Oxford: Elsevier. Volume 5, pp. 665-677.

Cunningham, H., Maynard, D., Bontcheva, K., and Tablan, V., 2002. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. In *Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02)*.

Gaunet, F., & Briffault, X., 2005. Exploring the functional specifications of a localized wayfinding verbal aid for blind pedestrians: Simple and structured urban areas. *Human-Computer Interaction*, 20(3), pp. 267-314.

Jackendoff, R., & Landau, B., 1992. Spatial Language and Spatial Cognition. In Jackendoff, R. ed. *Languages of the Mind: Essays on Mental Representation*. Massachusetts, USA: MIT Press. Ch. 6.

Kulyukin, V., Nicholson, J., Ross, D., Marston, J., & Gaunet, F., 2008. The blind leading the blind: toward collaborative online route information management by individuals with visual impairments. In AAAI *Proceedings from the 2008 AAAI Spring Symposium*. Stanford University, California 26 – 28 March 2008. AAAI Press.

Miele, J. A., Landau, S., & Gilden, D., 2006. Talking TMAP: automated generation of audio-tactile maps using Smith-Kettlewell's TMAP software. *British Journal of Visual Impairment*, 24(2), pp. 93-100.

Passini, R., & Proulx, G., 1988. Wayfinding without vision: An experiment with congenitally totally blind people. *Environment and Behavior*, 20(2), pp. 227-252.

Sendero Group LLC. 2009. Sendero GPS. [Online] http://www.senderogroup.com/shopgps.htm [Accessed 3 January 2009].

Talmy, L., 1983. How language structures space. In Pick, H. & Acredolo, L., eds. *Spatial Orientation*. New York, NY: Plenium Press. pp. 225-282.