REVIEWS

Review 1

Overall Evaluation: (0) borderline paper
Reviewer's confidence: (4) high

Review:
The paper addresses an interesting and topical issue of understanding place descriptions and extracting place expressions and representing them in a form that can be manipulated automatically.

The work relates to earlier work on spatial role labelling and seeks to extract "more informative" location expressions and describes some heuristics to do so. **The notion of informative is intuitive and described by examples. However, could it be more formalised to support the extraction process?**
- Thanks for your comment. We did realize that the term 'being informative' couldnt be conceptualized formally. From the mechanism of the approach, what can be said is that any co-occuring word which acts as a modifier to the place name or the spatial relation adds to the information. However, verb modifiers are regarded as exceptions to it. We feel that the intricacies of the notion could be stressed in an extended version.

I had trouble identifying the particular contribution of the work, where the relationship between the methods proposed and other components utilised from other sources (the Stanford parser, the tell us where corpus) was not presented clearly enough.
- Thanks for the comment. The reported work presents an effort to enhance the Stanford parser, so as to accommodate the extraction of spatial information from natural language place descriptions.
- The same line is now put at the start of second paragraph of Section 6, Conclusion to highlight the contribution.
- Last few lines in the implementation section are modified to clear the functionalities of Stanford parser.

A few examples from the application of the method were presented and discussed, but **no convincing systematic evaluation is presented of the quality of the resulting expression and particularly how it compares to the results of spatial role modelling approach.**
- We appreciate your concern, but considering the large set of descriptions, we could not set up an automated systematic means to compare the quality of our approach. We can only present samples in the form of examples and we tried to do that in the paper.
- We do provide a special mention for it at the start of Section 5 Experiment Study
- Considering the comparison with spatial role modelling. few lines added to indicate the differences:
1] Section 5.2 second paragraph "Nonetheless, from the output of the parser ..... a construction site.added to clarify the numbers relate to DLE identification and not triplet extraction

However, a good discussion of errors and complexities of the methods is presented.
Overall, it is interesting work (in progress) which I hope to read more about in the future.
Minor issues:
Some terminology is used without enough explanation, e.g. dialog-driven geolocation services and situated place description,
- Thanks for pointing out the terms for elaboration. We have now added few lines in the first paragraph of introduction
1]"To add to the complexity, the descriptions can be situated in a specific context and thus use indirect references (to its east, that building)."
2] query/dialog driven geolocation services (such as navigation, emergency assistance)

IOB encoding
- Few lines have been modified in the Implementation Section and the following line is introduced to simplify the description of an IOB encoding.
"IOB encoding is used .. in this case are the DLEs"
For clarification, the encoding holds no direct relevance to the paper so we chose not to elaborate it further owing to space limitations.

Page 2, section 3.1.1 - the word "none" in the set is not clear.
- Thank you again. The term has been removed.

==========

Review 2
Score: 1 weak accept
Reviewer's confidence: 4 high
The paper describes a method for identifying triples (although actually focusses only on the spatial relation and the reference object). The method was tested, and the paper describes a number of problems that resulted from the process.

The work is interesting, and the paper is well written and clear. I have two main concerns.
Firstly, in terms of the paper itself, there is no clear indication of the success after all the different errors are accounted for. I would think that there would be many of these kinds of errors, as detecting spatial from non-spatial expressions can be very challenging. **What were your actual success figures**? Of those that were identified as DLEs, how many were actually correct based on your manual checking? **Can you provide a measure of success - precision and recall?**
Thanks for the feedback. The change made is mentioned below:
- Table-4 added to provide the statistical results, the small description of which is provided in first para of Section 5.2

Is the 0.76 you mention at the end of Section 5.3.3 your overall measure of success without annotations?
- No, that's not the case. We appreciate the doubts raised and hence Footnote 6 has been added to clarify that the numbers are related to DLE identification and not triplet extraction .

It would also be useful to know a bit more about the annotations. Obviously they improve the results, but it really means the
method is not automated, which severely restricts the benefits?
Thanks for the your comments. We now mention under the first paragraph Section 5 that the the difference in the results are not significant when comparing the use of manual annotations.

My second main concern is that if **you extract only spatial relation and reference object, is that really enough to correctly interpret the expression spatially, without any other context?**
- We appreciate your concerns, however, for clarification, we don't just extract spatial relation and reference object, but a spatial triplet of the form locatum, spatial relation and the reference object

The more complex representations of language that you mention, like GUM-Space include a lot of other information to assist in correct interpretation, and even that it not really enough in my view. I think you plan to do some of the interpretation ahead of the extraction (e.g. you mention egocentric reference systems posing a challenge for extraction, and that you would translate into a relative reference system), but there are many other things to consider. To give a simple example, if you just extracted 'in the lake' or 'in Loch Ness', how would you know whether it meant in the 3D water geometry (there is a submarine in the lake); partially protruding (he is swimming in the lake); or 2D containment (there is an island in the lake) - there are other example: there's a pontoon in the lake. These kinds of ambiguities are widely discussed in the literature. **While it is very attractive and easy for computation to have a simple triple model, do you really think it can do the job when you come to try to identify what it means (and for example, unambiguously define the spatial geometry that is meant by 'in the lake' etc.**
- Very well pointed out as this indeed is a critical issue of sense disambiguation. It can be clarified that if carefully extracted, *island in the lake* **would be a valid triplet and** *swimming in the lake* **won't be. The idea is to learn the spatial sense of a preposition, which is a part of the future work. No modification were thus made to the paper in this regard.**

I also note that you ignore the verb, but this might also provide an important clue to the meaning of the expression. e.g. 'the river runs through the city' vs. 'the river meanders through the city'.

Minor points:

p.3 last paragraph before 'partial DLEs' title, **I would argue that the same reasoning applies for your example of we are at the Baretto's in the Alan.. Building - I think the second triple should be an in relation between the Baretto's and the building, not between we and the building.**
- Thank you for your analysis. However, there's difference of a verb and a preposition being used against two prepositions being used. In the mentioned paragraph, the case is of using a verb and a preposition. Examples of the format: "You <verb> PlaceX <spatial relation> PlaceY" should always yield triplet:
PlaceX <spatial relation> PlaceY for all verbs.

This is one of those famous ambiguities, in which you might get incorrect errors if you always assume the same interpretation, as I'm sure there are some examples that are context dependent. Actually, technically they should be distinguished by a comma ('the river flows through the park near the city' and 'the river flows through the park, near the city' - different meanings).

"I am 300 meters far from the auburn train-station" - this is a very unusual phrase in English, you would normally say 'I am 300 meters away from the auburn train-station' or colloquially 'I am 300 meters from the auburn train-station'
- Indeed it seems an unusual phrase, however these examples are directly taken from the datasets (tell us where/campus descriptions) and do contain some incorrect/informal usage. Since, there were more than handful number of such cases, we deal with them as it is.

'The problem can be defined similar to a spatial role labelling task' should be similarly (also just before section 4).
used for testing out methodology' should be our
- Thanks for pointing out the errors. Recommended modifications are now implemented.

'

You may also consider referring to the following, as they provide different conceptual models (e.g. Kracht provides a simple linguistic structure that has some similarities to yours):

Hornsby, K. and Li, N. (2009), Conceptual Framework for Modeling Dynamic Paths from Natural Language Expressions. Transactions in GIS, 13(s1), pp.27-45.

MacMahon, M., Stankiewicz, B. and Kuipers, B. (2006). Walk the Talk: Connecting Language, Knowledge, Action in Route Instructions, National Conference on Artificial Intelligence (AAAI-06).

Kracht, M. (2000). On the Semantics of Locatives. Linguistics and Philosophy, 25, pp.157-232.

Francez, N. and Steedman, M. (2006). Categorial grammar and the semantics of contextual prepositional phrases. Linguistics and Philosophy, 29(4), pp.381 - 417.

Mani, I., Hitzeman. J., Richer, J., Harris, D., Quimby, R., and Wellner, B. (2008). SpatialML: Annotation Scheme, Corpora, and Tools, In: Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odjik, Stelios Piperidis, Daniel Tapias (eds) Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco.

Mani, I., Doran, C., Harris, D., Hitzeman, J., Quimby, R., Richer, J., Wellner, B., Mardis, S. and Clancy, S. (2009). SpatialML: Annotation Scheme, Resources, and Evaluation. Technical Report.http://www.mitre.org/work/tech_papers/tech_papers_09/09_3827/

- Thank you for the suggested literature. We have read the papers and decided on adding those we felt were more appropriate for this workshop paper, while the rest we think will fit better in a more extended version of this work

==========

Review 3

Score: 2 accept

Reviewer's confidence: 3 medium

This paper describes an approach to extract spatial triplets describing qualitative spatial information from natural language text.

The authors present a very simple approach that combines two existing pieces of software: (1) a pre-trained, CRF-based extractor that identifies degenerate locative expressions (DLEs) (e.g. "in the Alan Gilbert Building"), and (3) the Stanford dependency parser. The authors basically combine the outputs of both components and use the Stanford parser to identify the subject of the DLE to generate a complete triplet.

The simplicity of the approach limits the contribution of the paper a bit because the key computations are done by the two existing components. This point is illustrated in the results discussion because the performance of the triplet extraction is basically dependent on the performance of the DLE extractor and the dependency parser. Nevertheless, the paper investigates a very relevant topic and the contribution and depth are appropriate for a workshop paper.

The authors state that they present a reasoning approach that is "devoid of any external resources (maps, path geometries or robotic vision)." To me, this is not entirely true because the DLE extractor uses a Gazetteer and/or manual annotations to identify place names. In addition, a manual annotation approach does not seem scalable.
- Thanks for the keen observation. We apologize for being not so elaborate in saying so. The use of external resources is totally optional and the tradeoff in the results is not significant. To stress on this, we have included the following lines
1] Last paragraph of Implementation section- "The use of manual annotations/gazeteers is optional, but can be included to enhance the DLE prediction scores" .
2] First paragraph of Section 5.2 "usage of manual annotations .. triplet extraction"

A strength of the paper is the detailed discussion of the results of the experiment using the Tell Us Where + Parkville campus description dataset. This gives the reader a good understanding of the strengths and weaknesses of the approach and clearly shows areas for future work. However, this section could be improved by giving quantitative measures for precision and recall over the entire dataset.

Some minor points:

In the introduction, the acronym DLE is used before it is defined, and the meaning of the term static relation was not clear to me in the context of this paper without some additional explanation
Thanks for the suggestion, in due of which following changes were made to the paper :
- Acronym removed.
- The term static relation removed everywhere while defining the term static triplet with an extra line added in the introduction of last paragraph "However for the purpose of extraction of triplets, we target only the static triplets which don't demand interpretation of spatial motion."

Major reductions in the paper to comply with the length of the paper:

1] Figure 2 for IOB encoding removed as the encoding holds no direct relevance to the paper.