Extracting Spatial Information From Place Descriptions

Arbaz Khan
Department of Computer Science and Engineering
Indian Institute of Technology Kanpur
Uttar Pradesh, India
arbazk@iitk.ac.in

Maria Vasardani
Department of Infrastructure Engineering
University of Melbourne
Victoria, 3010
mvasardani@unimelb.edu.au

Stephan Winter
Department of Infrastructure Engineering
University of Melbourne
Victoria, 3010
winter@unimelb.edu.au

ABSTRACT
A computational model of understanding place descriptions is a cardinal issue in multiple disciplines and provides critical applications especially in dialog-driven geolocation services. This research targets the automated extraction of spatial triplets to represent qualitative spatial relations between recognized places from natural language place descriptions via a simple class of locative expressions. We attempt to produce triplets, informative and *convenient* enough as a medium to convert verbal descriptions to graph representations of places and their relationships. We present a reasoning approach devoid of any external resources (maps, path geometries or robotic vision) for understanding place descriptions. We then apply our methodologies to situated place descriptions and study the results, its errors and implied future research.

Categories and Subject Descriptors
H.3.1 [**Information Storage and Retrieval**]: content analysis and indexing - *Linguistic processing*; I.2.7 [**Artificial Intelligence**]: Natural Language Processing - *Language parsing and understanding*

General Terms
Experimentation, algorithms, languages, performance

Keywords
Place descriptions, spatial role labelling, spatial language understanding, locative expressions, triplets

1. INTRODUCTION
A place description provides spatial information in terms of spatial features and the spatial relations between them. Place descriptions are used in everyday activities in spatial learning, problem solving and decision support in large-scale environments, independent and typically in absence of access to a map or a compass. The task of mapping a two-dimensional environment into a linear structure of language descriptions can also take a dynamic nature where a human uses a tour route for communicating spatial information [17, 5]. Thus, place descriptions can also be thought of encompassing path descriptions. To add to the complexity, the spatial features in the descriptions can be simple spatial features (house, park) or place names (official or informal). Nonetheless, this challenging task of understanding the natural language place descriptions provides critical applications in numerous domains including instruction following robots, description based localization in query/dialog based navigation services and automated geo-tagging of text. Also, having inferred the spatial relations from such a description, the translation of the relatively abstract verbal descriptions to more iconic media of conveying information (e.g., pictures and sketch maps) becomes easier.

This research targets the automated extraction of spatial triplets to represent spatial features and relations between them from natural language place descriptions via a degenerate form of locative expressions. *Locative expressions* as described by Herskovits [8], are expressions containing a preposition, its object and whatever the prepositional phrase modifies (i.e., subject of preposition). We, however, start from a degenerate form of locative expressions that are devoid of a subject. We provide a detailed definition for such expressions in Section 3.1.2. The challenge of this paper is to extract the subject of the DLEs to complete a spatial triplet.

The extraction process is accompanied with a trained learning model ([19]) which predicts the beginning and the span of a degenerate locative expression in a sentence using Conditional Random Fields (CRFs) along-with the place names indicators (gazetteers and manual annotations). It gives a good indication on the spatial relations in Natural Language (NL) text and can be used to extract triplets in the format of an object of interest, a reference object and a prepositional relation that connects the two . This representation is derived from the technique of Spatial Role Labelling (SpRL) [14], where the preposition serves as a spatial indicator connecting a trajector (object of interest) with a landmark (reference object). The end goal of this effort is to ease the translation of verbal descriptions to graph representations (with features as nodes and relations as edges). The aim is to make the triplets as *informative* and *convenient* as possible for such a translation. By being *informative*, we mean to exploit the spatial information in the three components of the triplets. Also, the triplet output format becomes *convenient* enough to a translator if it cuts down the linguistic complexities.

The introduction of the problem leads to the fundamental question: Given a degenerate locative expression, is it possible to extract the spatial triplets without introducing any additional ambiguities? If yes, how effective is it to use degenerate locative expressions (DLEs) to identify the underlying spatial relationships between the places. In this paper we present a technique to identify the spatial relationships between recognized objects and place names with the help of DLEs. We claim our identified spatial relations to be qualitative and informative enough to assist in sketch-map generation and graphical depictions. However for the purpose of extraction of triplets, we are targeting only the static relations in the descriptions. We then investigate the different types of errors in the results of triplet extraction and highlight the gaps in the approach of its inability to resolve the indirect reference in a description and end by pointing out the challenges for dealing with full-blown English place descriptions.

2. RELEVANT WORK

We are interested in the extraction of geographic information in the form of triplets, from unrestricted NL place descriptions. Research in the past had focused on mining specific and application-dependent geographic information from controlled language expressions [9, 16, 22] or from certain contexts, such as route descriptions. Linguistic constructs are often complex and lead to under or over-specifications when trying to match NL to formal frameworks of spatial information. As argued in [4], most computational aspects of geographic information do not consider these linguistic intricacies, especially in the construction of formal models of spatial relations based on the logic of human spatial cognition. The authors also argue that such formal models do not acknowledge the way that people actually express spatial relations linguistically. Therefore, Bateman [3] suggests a semantics and linguistics based spatial ontology that would facilitate better mappings between spatial calculi and NL spatial expressions, such as the Generalized Upper Model ontology (GUM) [2].

In more recent works, researchers have been working on parsers that can process unrestricted language input, and thus, come closer to the parser discussed in this paper. For example, Zlatev [26] worked on the cognitive linguistic research in spatial semantics and provided basic theoretical concepts for understanding it, in the form of the holistic spatial semantic theory. Based on that theory, Kordkamshidi et al. [13] then implemented spatial role labeling (SpRL), a method with which they assigned reference objects, locata and spatial relations roles to linguistic terms. This technique facilitates mapping the terms onto formal spatial relations and is also used in this work. The authors suggest machine learning techniques to deal with ambiguity in linguistic spatial information and in [14], they utilize the proposition project (TPP) from SemEval-2007 [18] for disambiguating the spatial meanings of prepositions and enhancing the SpRL technique. Zlatev [26] worked on the cognitive linguistic research in spatial semantics and provided basic theoretical concepts for understanding it. Ever since then, there has been growing trend to understand natural language (NL) descriptions. The research has moved towards understanding unrestricted NL in current works ([14, 20, 11]). The methodology suggested in this paper makes use of no external resources (such as maps, vision), but employs a machine learning tool [19] to identify place names and the associated prepositional phrases. Hereafter, we work with the concept and reasoning approach to form rules for extracting useful spatial information from the text.

3. THEORY

3.1 Definitions

In this section, comprehensive definitions of the major terms that appear in this paper, viz spatial triplets (or simply triplets) and degenerate locative expressions (DLE), are provided.

3.1.1 Spatial Triplets

A spatial triplet is a triple of a locatum (LOC), reference object (RO) and a spatial relation (r). Here, RO is the object by which the location of the LOC is defined, using the prepositional relation r. For example, `<Melbourne Hospital` *diagonally across* `Peter Doherty Institute>` references *Melbourne Hospital* as the LOC in terms of *Peter Doherty Institute* as RO by the relation *diagonally across* as r. Vasardani et al. [24] provide more information on the theory and the structure of triplets. This method of defining spatial relation is derived from the technique of Spatial Role Labelling(SpRL), which first appeared in [14]. In SpRL, one assigns spatial role labels to the words or phrases in sentences from the set {*trajector* (tr), *spatial Indicator*(si), *landmark*(l), *none*}. For example, such a labelled sentence looks like: "$[We]_{tr}$ are sitting $[in]_{si}$ $[\textit{the Baretto's}]_l$ $[in]_{si}$ $[\textit{the Alan Gilbert Building}]_l$". However, spatial triplets that could be extracted out of such a sentence look like :
`<We` *in* `the Baretto's>`
`<We` *in* `the Alan Gilbert Building>`
Arguably, a spatial triplet relation provides a more convenient representation of spatial information than the role labels, and provides a suitable medium to translate the verbal description to a graph representation.

3.1.2 Degenerate Locative Expressions (DLE)

A locative expression is a spatial referring expression which contains a preposition, its object, and whatever the prepositional phrase modifies (i.e., subject of preposition) [8]. These expressions have been adopted as a standard for spatial semantics and cognition and are widely referred to [21, 26]. However, in informal communication or a situated dialog, people tend to disregard the subject and describe the location in terms of just the preposition and its object (e.g., [I am] at Deakin University). We term such locative expressions as degenerate locative expresssions (DLE). In a crowd-sourced corpus of place descriptions[1] people described their location mostly in terms of DLEs (e.g., *at the University of Melbourne*, *in my house*). Hence, an attempt was made to identify the DLEs and train a model on this available dataset to provide predictions on any test corpus.

[1]Data collected by a mobile location-based game, http://telluswhere.net/

Classification of Degenerate Locative Expressions

From the definition of a degenerate locative expression (DLE), one can think of constructions like "of the train-station", "from the University" and "to Bourke Street". Such examples do follow from the definition but cannot convey any spatial relationship independently. Binary spatial relationships are not informative when naively using non-locative prepositions (e.g., X *from* Y is not an informative relationship). Such DLEs require special treatment before the corresponding triplet relation could be extracted. We term such DLEs as *partial DLE*s. These expressions include all those spatial expressions which do not start with locative prepositions[2]. The rest of the DLEs, which start with a preposition of location and thus give a sense of spatial relation are termed as *locative DLE*s. For example, *in the alley* is a locative DLE. The procedures for triplet extraction described in the subsequent sections are described first for locative DLEs and then for partial DLEs.

3.2 Extracting Spatial Triplets

To deal with the task of spatial information extraction, we set ourselves the aim of identifying 'informative' triplets. The problem can be defined similar to a spatial role labelling task (SpRL), the difference being that we attempt not to miss any information from the place names and the spatial relation. For an example from the SpRL dataset, "two men are standing on a lawn in front of a light brown house with blue corners", SpRL marks one of the reference object as 'house' to eventually give the triplet <lawn *in front of* house>. However, we are interested in including any informative meaning inherent in the description and thus output "lawn in front of light brown house". This also becomes essential in the long term goal of translating verbal descriptions to graphical depictions, where the graph/sketch drawing algorithm could be vitally benefitted by localizing the spatial feature to 'light brown house' rather than just any 'house' (from the above example). There have been various datasets used for identifying spatial expressions [2, 7, 12] but none are characteristic enough of a place description. For instance, place descriptions are more descriptive than mere spatial expressions and the spatial relations can be fairly complex and difficult to formalise (e.g., 'between', 'across from'). Due to the goals set thereafter and the fact that there is no available dataset for training on such special and essential requirements, a reasoning approach, rather than a machine-learning approach is chosen here for extracting triplet relations. This approach helps in understanding place descriptions and thus eventually aiding machine learners by revealing the structure and identifying relevant/irrelevant attributes.

Extracting triplet relations poses varying challenges owing to the complex nature of place descriptions. The spatial relations can be defined either simply using position descriptors (e.g., on, at, in) or using concepts of motion and indirect references [26]. For example, in "From here, you can go left towads X and then head north to reach Y", the phrase "from here" carries context (indirect reference) and the current context is shown spatially related to X and Y via motion indicators. For the purpose of this paper we limit ourselves from the detection of spatial motion and resolution of indirect references and instead focus on the structure of static place-descriptions. The triplet extraction assumes all descriptions are static, i.e., without any context or motion. This process of extraction of triplets goes through the DLEs by categorising them into the two types: *locative* and *partial*. We start by describing the process of triplet extraction for locative DLEs, and then move forward to deal with partial DLEs.

Locative DLEs

Given the DLEs, to identify the triplet relation, one needs to find the subject of DLE. For example, in "There is a construction site in the alley", the DLE is 'in the alley' and the triplet relation extracted should be <construction site *in* the alley>. To identify the subject, the dependency between the preposition *in* and the two noun places 'construction site' and 'the alley' need to be identified. Luckily, from the Stanford typed dependencies parser [6] one can extract the dependencies between the words of a sentence. Using the collapsed dependencies representation, *prep* relation reports a prepositional phrase and the verb/noun/adjective modified by the prepositional phrase. In the syntactic terminology of the Stanford dependency parser, the former is called the dependent and the latter is called the governor. In this case, the dependency reported is *prep_in*(site,alley) where 'site' is the governor and 'alley' is the dependent. When the relation is between two nouns, the triplet relation is direct. However, in the case of verb/adjective, we further explore the typed dependencies to find the subject or the noun modified by the corresponding verb/adjective. The subject finding task (for verb/adjective/noun) then recursively checks for the modifiers of the subject (if any) by searching through the set of modifier relations provided by typed dependencies. The final triplet relation is reported thereafter. The steps to extract triplets from a sample sentence for the case of *prep* relation involving a verb is shown in Figure 1.

In descriptions where a person describes a location using references to pronouns, e.g., *You can also find NAB and Commonwealth bank branches near the Union House*, the direct object of the verb gets linked to its prepositional relation (*near*). So, in this sentence, subject of the verb *find* is *You*, but the direct objects are *NAB* and *Commonwealth Bank*, and thus the triplet should be <NAB *near* Union House> (and not *You near Union House*). Thus, the heuristic that works here is: if there is a direct object of the verb, report the direct object as the locatum of the triplet. The heuristic can be further strengthened by using the fact that in such cases the nominal subject of the verb is a pronoun.

Partial DLEs

The partial DLEs convey no spatial information independently but they form a medium for inferring spatial relationships in the place description. As mentioned earlier, partial DLEs can be converted to locative DLEs in some sentences and can be directly used to output triplets. In other cases, a partial DLE can be made informative and meaningful by just extending its prepositional clause. For example, consider the sentence "I am 300 meters far from the auburn train-station" with the partial DLE "from the auburn train-

[2]Here we don't deal with the exhaustive list of locative prepositions (Jackendoff and Landau [15]) but focus only on *at, in* and *on* as locative prepositions, the rest of the prepositions are treated as non-locative.

Table 1: Extending the prepositional relations of direction using Stanford dependencies (DEP)

Governor	Action on relation	Example	
		Sentence	**Actions**
noun	extend with the noun and its modifiers	Its the 3rd house from the station.	DEP : **prep_from**(house, station) extends *from* to "3rd house from"
adjective	extend with the adjective	My house is next to the station.	DEP : **prep_to**(next, station) extends *to* to "next to"
verb	extend with the verb's modifiers*	I am living 2 minutes from the station.	DEP : **prep_from**(living, station) adds modifiers of *living* ('2 minutes') extends *from* to "2 minutes from"

*-preference to verb's direct object (if any)

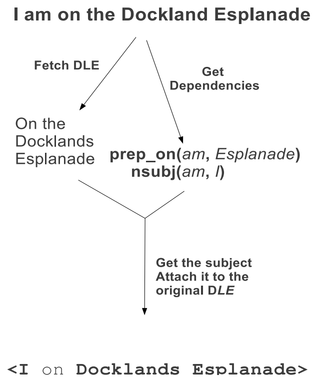

Figure 1: Sample run of the identification process for locative DLE in the sentence "I am on the Docklands Esplanade"

station". Clearly, by including the "300 meters far [from]" we get a more informative triplet `<I 300` *meters far from* `the train station>`. Thus, the aim is to capture this information from the typed dependencies. Table 1 provides the exhaustive rules to extract this information from the typed dependencies. By doing this, we exploit the locative purpose served by a partial DLE. Thereafter, we proceed similar to a locative DLE to find subject of the preposition.

4. IMPLEMENTATION

The methodology mentioned in the previous section works on the natural language via the results of a parser which is trained to identify degenerate locative expressions in informal text [19]. The parser builds a *conditional random field* (CRF) machine learning model on the manually annotated corpus of place descriptions [1] and attempts to identify both informal and formal place references. To enhance its prediction scores, it exploits external resources such as gazetteers and dictionaries. The output provided by the parser is in the form of IOB encoding as marked in Figure 2. A B-NP tag denotes the *beginning* of DLE, while I-NP tags indicate the word is *in* the span of a DLE. All words with O tags

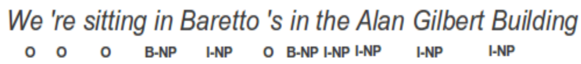

Figure 2: IOB-encoded identification of degenerate locative expressions

are *outside* of any DLE. Hence, in the example shown in Figure 2, "in Baretto" and "in Alan Gilbert Building" are identified as DLEs in the sentence. The DLEs output by the parser are either simply place names (e.g., Chetwynd Place) or place names with prepositional attachments (e.g., of the train station).

The parser is written in *Python* and our program works on top of it. The input requirements for the parser are met by *Part-of-Speech* (POS) tagging and shallow parsing (*chunking*) the sentences. POS-tagging corresponds to determining the part-of-speech tag (e.g., noun, verb, adverb) associated with a word, whereas shallow parsing is the process of grouping words in a sentence as *chunks* to analyse the syntactic structures. Besides this, the parser can be aided in locating place references by marking the place names and manually annotating them with information such as granularity and identifiability. The format of these manual annotations follows the scheme provided by Tytyk and Baldwin [23]. The manual annotations are optional, but can be included to enhance the DLE prediction scores. For each DLE extracted from the parser, our program identifies its subject (after extending prepositional clauses for partial DLEs) using the output of the *Stanford parser* [10] for POS tags and dependencies in the original sentences.

5. EXPERIMENTAL STUDY

In this section, the experimental analysis of our methodology is presented. First, the dataset used for testing out methodology is described. Next, an insight to the performance of our approach for extracting triplets from our corpora, the *Tell Us Where* mobile game dataset and the set of campus descriptions, is given. Finally, the errors produced in extracting the triplets and the sources behind them, as well as the types of triplets missed by the approach are discussed.

5.1 Dataset Description

For our experimental study, we used the *Tell Us Where* dataset [1] as one of the test corpora. The dataset comes from a location based mobile game which locates its users and asks them: "Tell us where you are". Originally, the corpus consisted of 1,858 place descriptions but for the purpose of this research, descriptions containing fewer than 20 words are filtered out. This gives 54 distinct situated place descriptions, each of which describes relationships among two or more places. Additionally, a set of four NL descriptions of the Parkville campus of the University of Melbourne is parsed, none of which originally are less than 100 words. These were submitted by graduate students with varying degrees of familiarity with the campus. Also, there were no instructions given to the students while submitting the descriptions, which yields a set of unrestricted NL place descriptions. With such datasets, our testing model is set up to be descriptive enough, constraint-free and thus more natural in relation to, and characteristic of place descriptions.

5.2 Extraction Of Triplets

The task of extracting 'spatially informative' triplets from NL descriptions is challenging and cannot be directly dealt using simply a language parser. The DLE parser aids in separating the place names, and manual annotations to enhance the DLE parser are employed[3]. This subsection focusses on the triplets extraction study pertaining to static place descriptions. The parser's output in such cases has a fair similarity with the expected triplets. Below are some of the example cases representative of the results.

EXAMPLE 1. *I am* [*at a 3 stories town house*]*, number 7* [*Chetwynd Place*]. [*The house*] *is located* [*in a small alley*] [*behind a row of town house*] [*along Chetwynd Street*]*,* [*near the corner*] [*between Chetwynd Street and Queensberry Street*]*. There is* [*a construction site*] [*in the alley*]*. Its the 3rd house from the head* [*of the alley*]*.*

This example is taken from the *Tell Us Where* corpus. The bracketed texts correspond to the DLEs identified by the parser. The triplets output by the program are presented in Table 2. This is a characteristic example of a static place description without any indirect reference, except in the last sentence where 'it' is used in place of "the house" and thus no triplet was extracted. If the sentence is rephrased as "*I am in the 3rd house from the head of the alley*", the triplet extracted is, `<I` *in 3rd house from* `the head of the alley>`[4]. Not all triplets inherent in the description were extracted (see Section 5.4). Nonetheless, from the output of the parser, the triplets produced seem promising and interpretable for an algorithm for graphical depiction.

The 'between' relation in the example is spatially informative, but not syntactically suitable for a triplet, as it has two reference objects connected to the same locatum. This inhibits the computational interpretation of triplets. Vasardani et al. [24] discuss the semantics of triplet relations, and propose that triples with 'between' relation can be broken into two, by identifying the conjuncts of the dependent (here, 'Chetwynd Street' and 'Queensberry Street'). However, the conjuncts were not separately identified in this case because of the presence of the connector 'and', which forces the DLE parser to identify *Chetwynd Street and Queensberry Street* as one single place name, thus leading to a bias in the subject finding methodology. In a different case, where such a restriction does not hold because the word 'the' follows a connector 'and', triplets produced for a *between* relation are shown below (Example 2). The output comes out as suggested in [24].

`<I` *at* `a 3 stories town house>`
`<The house` *in* `a small alley>`
`<a small alley` *behind* `a row of town house>`
`<The house` *along* `Chetwynd Street>`
`<The house` *near* `the corner>`
`<the corner` *between* `Chetwynd Street and Queensberry Street>`
`<a construction site` *in* `the alley`

Table 2: Parser output for spatial triplet relations in Example 1

EXAMPLE 2. *Between this building and* [*the campus*]*,* [*on Swanston St*]*, is the major public transport hub for* [*the University*]*.*

Triplets extracted:
`<the major public transport hub` *between* `building>`
`<the major public transport hub` *between* `the campus>`
`<the major public transport hub` *on* `Swanston St>`

EXAMPLE 3. *We're sitting* [*in Baretto*]*'s* [*in the Alan Gilbert Building*]*,* [*across Grattan street*] *is one of the* [*medical buildings*]*.* [*Down the hill*] [*along Grattan street*] *the* [*new building*] *being constructed is the* [*Peter Doherty Institute*] *and diagonally across the road (*[*Royal Parade*]*) is* [*Melbourne Hospital*]*. In the other direction the open area is* [*University square*] *(there is a* [*carpark*] *underneath). At the city end* [*of University square*] *is* [*the law building*]*.* [*Across Grattan street*] [*from University square*] *there is* [*an entrance*] [*to the campus*]*, straight ahead is* [*an overpass building*] *and to the right are the various* [*Engineering buildings*]*. The road goes in a big loop around* [*the campus*] *you can either go left* [*towards the Medical buildings*] *or right passed* [*the Engineering buildings*] *then head North (away from University square). One other area you may want to explore is* [*South Lawn*] *which you can get to by going* [*underneath the overpass building*] *directly in front of you when you enter* [*the campus*]*.*

This is an example taken from the set of campus descriptions. The triplets output by the program are shown in Table 3[5]. We observe that the static spatial relations come

[3]The results do not differ considerably with the set-up of not using the manual annotations except for the errors described in 5.3.3

[4]See how the preposition 'from' gets extended to '3rd house from', and the clause related to the preposition 'of' gets attached into the reference object as 'head of the alley'

[5]The first and the third sentence were split at clause boundaries to eliminate Stanford parser errors; see Section 5.3.2

<We *in* Baretto>
<We *in* Alan Gilbert Building>
<one of medical buildings *across* Grattan street>
<the Peter Doherty Institute *Down* the hill>
<the hill *along* Grattan street>
<the law building *at the city end of* University square>
<an entrance *across* Grattan street from University square>

Table 3: Program output for spatial triplet relations in Example 2

up as desired. For instance, the program infers the complex triplet relation:

`<an entrance` *`across`* `Grattan Street` *`from`* `University Square>` with its complete inherent information using the DLEs related to it, i.e., [Across Grattan street], [from University square] and [an entrance]. Again for formalising reasons, the 'across from' relation can be broken analogously to a 'between' relation in the previous example. It can be seen, however, that not all triplets were extracted by the program because of issues such as indirect references in the phrases "*straight ahead . . . Engineering buildings*" and "*In the other direction . . .*", and spatial motion in the sentence, "*The road goes . . .*", which are not dealt with currently by the parser.

5.3 Investigating Errors

From the results obtained by the test corpora, there was a clear indication that the validity of the produced triplets relies on the correctness of the Stanford typed dependencies, as well as the accuracy of the DLE parser. Some invalid cases are also the product of the subject finding methodology for the DLEs, mainly due to its inability to disambiguate the spatial sense in the subject. In other cases, it is the informal or incorrect usage of grammar that led to the failure of triplet extraction. Each of these cases are discussed in detail below.

5.3.1 Invalid Subjects

The errors corresponding to invalid subjects found can be prominently classified into two types: *invalid place name* and *subjects with unwanted information*. Since there is no spatial sense disambiguation done for the noun phrases, some of the subjects reported turn out to be invalid. For instance, in example 5, the subject of the DLE *along the street* is *approximately halfway* (a noun phrase) and is deemed correct by the program, producing an incorrect triplet. Similarly in example 7, subject of *at the gate and the house* is identified as *roses* (instead of *an arch*), which is not a place reference and thus is invalid. Such errors can be corrected by choosing subjects, which are identified as place references by the DLE parser. However, this leads to the side-effect of making the production of triplets heavily dependent on the accuracy of DLE parser.

Additionally, this approach of finding the subject of a DLE in the most informative way leads sometimes to extensions with undesired information. In the example 4, *Wilson Hall* gets undesirably extended to *Wilson Hall a multi-function hall*, due to having DLEs separated by a comma identified as a single expression. While in most cases this rule provides useful spatial information (e.g., the University of Melbourne, in Carlton), in the example provided here the expression after the comma does not add to the spatial sense of the DLE.

Example 4. *Next* [*to the sandstone core*] *is* [*Wilson Hall*]*, a multi-function hall . . .*

Triplet extracted:
`<Wilson Hall a multi-function hall` *`next to`* `the sandstone core>`

5.3.2 Limitations Of The Stanford Parser

The errors corresponding to limitations of the Stanford parser include incorrect POS tagging and hence, incorrect typed dependencies. Since, the backbone of the subject finding task are the typed dependencies, the major errors that occur in the extraction process are when the dependencies are incorrect. For instance, in example 5, the parser failed to realise the dependency between *private residence* and *along the street*. Some of the errors originate from the incorrect POS tags such as that in Example 3, where *diagonally* is identified as a noun(instead of an adverb), and thus an invalid triplet is produced.

Example 5. *I am* [*at a private residence*] *located on the western side* [*of Barrington Avenue*] [*in Kew*]*, approximately halfway* [*along the street*]*.*

Triplets extracted:
`<I` *`at`* `a private residence>`
`<a private residence` *`on the western side of`* `Barrington Avenue>`
`<Barrington Avenue` *`in`* `Kew>`
`<approximately halfway` *`along`* `the street>`

The study of the error of these types led to an interesting observation. The major cases where the Stanford parser fails to identify the dependencies are those with long sentences containing multiple clauses. This can be fairly justified considering the inexhaustive training behind the Stanford parser. However, it can be seen that such long sentences have visible clause boundaries (usually separated by ',') and if the clause boundaries are detected and the sentence is split into its clauses, the dependencies turn out correct. Clause separation was used whenever possible in the triplet extraction to eliminate dependency errors. Correcting the Stanford parser is clearly out of scope for this paper. However, for future uses, it is suggested to either retrain the parser to handle multi-clausal sentences, or alternatively, to preprocess the data by splitting the sentences at their clause boundaries, before feeding them to the parser.

5.3.3 Inaccuracy Of The DLE Parser

Since the task of DLE parser is to identify prepositions and the following place references, errors are related to the reported span of the DLE and the preposition's spatial sense. Example 6 is a case where the lack of disambiguation of the preposition's sense is exposed. The DLE parser identifies *at night* as a DLE although *at* here has no spatial sense, thus resulting in an invalid triplet.

EXAMPLE 6. *There is* [*constant construction*] *which keeps all residents up* [*at night*].

Triplet extracted:
`<constant construction` *at* `night>`

Next example (Example 7) highlights the case where an incorrect span of a DLE is reported. The triplet extracted is `<roses` *at* `the gate and the house>` instead of `<roses` *at* `the gate>`, because of the effect of the connector 'and' mentioned previously (Section 5.2).

EXAMPLE 7. *There is an* [*arch*] *with roses growing on it* [*at the gate and the house*] *is a double -* [*fronted Victorian house*].

As compared to the F-score of the DLE parser without the manual annotations (0.76), the F-score of the parser with annotations is very high (0.99) [19]. Though using manual annotations increases the overall F-score of the DLE identification, it is observed in some descriptions the recall increases when not using the manual annotations, at the expense of accuracy. With the manual annotations, the parser becomes strictly restricted to the annotated place references and a DLE gets identified only if the associated place reference is present in the annotations. But for large enough datasets of place descriptions, manual annotations can miss out marking general place references such as house, street, or wall.

5.3.4 Incorrect Grammar Usage

The informal or incorrect usage of grammar may become a major obstruction in the extraction of useful spatial information. This also includes the concatenation of sentences without appropriate punctuation. Unfortunately in such cases, neither the Stanford parser, nor the DLE parser is robust enough to avoid the errors.

5.4 Missing Triplets

The results of triplet extraction have a good precision in static place descriptions containing no indirect or exophoric place references, or spatial motion. But for the corpus as a whole, besides dependencies missed by the Stanford parser, the inability of the approach to handle the limiting cases results in a low recall in terms of total expected triplets. For example, in Example 3, though about 80% of static relations were extracted, only about 40% of the total expected triplets (i.e., including those representing dynamic relations and indirect references) were reported. This outcome is understandable as the parser has no access to motion indicators and no specific approach targeting reference resolution.

5.5 Inability To Resolve Indirect References

One other obstacle found in the performance of our approach was its inability to resolve references. In a situated place description, people frequently make use of indirect references. The indirect references can be exophoric as well as implicit. For example, "*Near this cornering point, you occasionally find a crepes stand*" makes use of exophoric reference, but "*To the West is the Baillieu Library*" makes use of an implicit reference. Thus, the task of resolving indirect references subsumes the task of exophoric reference resolution and the previous well-studied domain to resolve exophoric references can not deal completely with these indirect references. However, it was observed that the indirect references used in human descriptions go in a depth-first fashion especially to describe descriptions involving path intersections. For example, in the description , "*there are three main alternative paths that you can use to head north. The central one is up some stairs* [...]. *The path will take you to the* [...] *where there is* [...] *and a little to the east, the Union House. Union House is a large building containing* [...]". And then the reference ends and the speaker switches back to the other branch to explore a new depth, "*The second path* [...] *from the south entrance takes you a little bit to the west and then switches north* [...]. *Near this* [...], *you find a crepes stand. From this path* [...]" (here expressions like "this path", "the road" represent the current reference). And then again it switches back, "*The third path* [...]". However, if there is a single path to describe, it comes out as a tree with a single branch. Though intuitively simple to understand, the complexity behind this approach is to infer the decision of whether moving further on the depth of the tree or to add a new branch to the existing depth.

6. CONCLUSIONS AND FUTURE WORK

This paper addresses the task of extracting spatial information from NL place descriptions by using triplet representations. We have come up with a simple, yet effective reasoning approach for understanding static place descriptions, without making use of any external resources such as maps, or path geometry. The attempt to find informative and computationally convenient triplets sets up a good potential medium for translating textual place descriptions to graph representations, such as automatically produced sketch maps, as suggested elsewhere [24]. We have implemented the approach to two different corpuses, which were not mere set of spatial expressions, but situated place descriptions. Our experiments investigated the applicability of the approach and indicate a good recall for extracting static spatial relationships for the production of triplets of a locatum, spatial relations and reference objects. It suggests that when the place descriptions contain a static structure in the spatial relationships, we can be confident of the reasoning approach to yield the relations, given that the DLE parser has identified the DLEs correctly. We have made an extensive study into the shortcomings of the approach exposing the impact of absence of indirect reference resolution and spatial motion interpretation in understanding situated place descriptions.

This research also forms the first attempt to use DLEs as an intermediate step to extract triplets from spatial language. Our experiments tested the DLE parser on a corpus different from the one it was orginally trained for. And, our analysis on the results of DLE parser suggests that use of the DLE parser is well suited for understanding static place descriptions without any indirect place references. We also pose the need for learning the prepositional sense to improve the precision of such a parser meant to identify DLEs. Further-more, we reveal that the Stanford dependency parser when used with place descriptions, errs in fundamental tasks of *POS*-tagging and extracting *typed dependencies*, and that the parser could achieve more reliable results if it is retrained to work on sentences with multiple clauses. And it remains as a critical tool to explore spatial language further by aiming at descriptions involving motion.

Our attempt to understand place descriptions exposes several gaps in this research area. One obvious direction of extending the reasoning approach in this research and making it more robust is to migrate to a case based reasoning (CBR) approach [25]. CBR is a learning approach that works by forming generalizations of training examples, leading to a more natural mechanism for triplets extraction. As highlighted earlier, resolving indirect references in a place description is a challenging task. References may point to place names and are crucial to resolve for extracting the spatial relationships. Such references can occur using indicators such as 'here', 'this', 'there','that' or 'it'. In other non trivial cases, there may be no explicit indication of the reference, but an implicit one in the context of the description. For instance, in "*In the center are the original 'old' buildings.*", there is no way one can resolve the indirect reference without identifying the context from the corresponding situated description. Hence, there rises a need of an approach which keeps track of the context and enables grounding of such non-trivial references. Additionally, extracting spatial relationships from a description becomes more challenging with an egocentric frame of reference, e.g., "*Facing the activity center, you would find a tower on your left*". One needs to figure out the relationship between *activity center* and the *tower* using the orientation set in the description and develop a generalized approach to deal with an egocentric frame of reference.

7. REFERENCES

[1] *Tell Us Where*. http://telluswhere.net. Location based mobile game.

[2] J. Bateman, T. Tenbrink, and S. Farrar. The role of conceptual and linguistic ontologies in interpreting spatial discourse. *Discourse Processes*, 44(3):175–212, 2007.

[3] J. A. Bateman. Language and space: A two-level semantic approach based onÂăprinciples of ontological engineering. *International Journal of Speech Technology*, 13(1):29–48, 2010.

[4] J. A. Bateman, J. Hois, R. Ross, and T. Tenbrink. A linguistic ontology of space for natural language processing. *Artificial Intelligence*, 174(14):1027–1071, 2010.

[5] M.-P. Daniel, L. Carite, and M. Denis. Modes of linearization in the description of spatial configurations. In J. Portugali, editor, *The Construction of Cognitive Maps*, volume 32 of *GeoJournal Library*, pages 297–318. Kluwer Academic, Dordrecht, 1996.

[6] M.-C. de Marneffe and C. D. Manning. Stanford typed dependencies manual. http://nlp.stanford.edu/software/dependencies_manual.pdf, 2008.

[7] M. Grubinger, P. Clough, H. Müller, and T. Deselaers. The IAPR TC-12 benchmark: A new evaluation resource for visual information systems. In *Proceedings of the International Workshop OntoImage 2006 – Language Resources for Content Based Image Retrieval*, pages 13–23, 2006.

[8] A. Herskovits. Semantics and pragmatics of locative expressions. *Cognitive Science*, 9(3):341–378, 1985.

[9] J. D. Kelleher. *A perceptually based computational framework for the interpretation of spatial language*. PhD thesis, Dublin City University, 2003.

[10] D. Klein and C. D. Manning. Accurate unlexicalized parsing. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*, pages 423–430. Association for Computational Linguistics, 2003.

[11] T. Kollar, S. Tellex, D. Roy, and N. Roy. Toward understanding natural language directions. In *5th ACM/IEEE International Conference on Human-Robot Interaction*, pages 259–266, 2010.

[12] P. Kordjamshidi, S. Bethard, and M.-F. Moens. SemEval-2012 Task 3: Spatial role labeling. In *Proceedings of the SemEval-2012 Evaluation Exercises*. Association for Computational Linguistics, 2012.

[13] P. Kordjamshidi, M. Van Otterlo, and M.-F. Moens. From language towards formal spatial calculi. In *Workshop on Computational Models of Spatial Language Interpretation (CoSLI 2010)*, 2010.

[14] P. Kordjamshidi, M. Van Otterlo, and M.-F. Moens. Spatial role labeling: Towards extraction of spatial relations from natural language. *ACM Transactions on Speech and Language Processing*, 8(3):4:1–4:36, 2011.

[15] B. Landau and R. Jackendoff. "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16:217–217, 1993.

[16] H. Li, T. Zhao, S. Li, and J. Zhao. The extraction of trajectories from real texts based on linear classification. In J. Nivre, H.-J. Kaalep, K. Muischnek, and M. Koit, editors, *Proceedings of the 16th Nordic Conference of Computational Linguistics NODALIDA-2007*, pages 121–127, University of Tartu, Finland, 2007.

[17] C. Linde and W. Labov. Spatial networks as a site for the study of language and thought. *Language*, 51(4):924–939, 1975.

[18] K. Litkowski and O. Hargraves. SemEval-2007 Task 06: Word-sense disambiguation of prepositions. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 24–29. Association for Computational Linguistics, 2007.

[19] F. Liu. Automatic identification of locative expressions from informal text. Master's thesis, University of Melbourne, Australia, 2013.

[20] C. Matuszek, D. Fox, and K. Koscher. Following directions using statistical machine translation. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, pages 251–258. IEEE Press, 2010.

[21] P. Olivier and J.-i. Tsujii. A computational view of the cognitive semantics of spatial prepositions. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pages 303–309. Association for Computational Linguistics,

1994.

[22] D. A. Tappan. *Knowledge-based spatial reasoning for automated scene generation from text descriptions.* PhD thesis, New Mexico State University, Las Cruces, NM, USA, 2004.

[23] I. Tytyk and T. Baldwin. Component-wise annotation and analysis of informal place descriptions. In M. Vasardani, S. Winter, K.-F. Richter, K. Janowicz, and W. Mackaness, editors, *International Workshop on Place-Related Knowledge Acquisition Research*, volume 881 of *CEUR*, pages 7–12, 2012.

[24] M. Vasardani, S. Timpf, S. Winter, and M. Tomko. From descriptions to depictions: a conceptual framework. In *Spatial Information Theory*, Lecture Notes in Computer Science, Berlin, 2013. Springer.

[25] L. D. Xu. Case based reasoning. *IEEE Potentials*, 13(5):10–13, 1994.

[26] J. Zlatev. Spatial semantics. *Handbook of Cognitive Linguistics*, pages 318–350, 2007.