UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

JOSÉ LUCAS SILVA FREITAS

TÍTULO: LOCATING PLACES DESCRIBED IN NATURAL LANGUAGE USING HUMAN-MACHINE DIALOGUE

CAMPINA GRANDE - PB

2021

F866l Freitas, José Lucas Silva.
Locating places described in natural language using human-machine dialogues / José Lucas Silva Freitas. - Campina Grande, 2021.
96 f. : il. Color

Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática, 2021.
"Orientação: Prof. Dr. Cláudio Elízio Calazans Campelo".
Referências.

1. Linguagens Conversacionais. 2. Modelagem Conceitual. 3. Processamento de Linguagem Natural. 4. Sistemas de Informação Geográficos. 5. Sistemas Converscionais. I. Campelo, Cláudio Elízio Calazans. II. Título.

CDU 004.432.45(043)

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECÁRIA ITAPUANA SOARES DIAS CRB-15/93

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
POS-GRADUACAO CIENCIAS DA COMPUTACAO
Rua Aprígio Veloso, 882, - Bairro Universitário, Campina Grande/PB, CEP 58429-900

FOLHA DE ASSINATURA PARA TESES E DISSERTAÇÕES

JOSÉ LUCAS SILVA FREITAS

LOCATING PLACES DESCRIBED IN NATURAL LANGUAGE USING HUMAN-MACHINE DIALOGUE

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação como pré-requisito para obtenção do título de Mestre em Ciência da Computação.

Aprovada em: 28/01/2021

Prof. Dr. CLÁUDIO ELÍZIO CALAZANS CAMPELO, UFCG, Orientador

Prof. Dr. CARLOS EDUARDO SANTOS PIRES, UFCG, Examinador Interno

Prof. Dr. YURI ALMEIDA LACERDA, IFCE, Examinador Externo

sei! assinatura eletrônica

Documento assinado eletronicamente por **Yuri Almeida Lacerda**, **Usuário Externo**, em 28/01/2021, às 16:16, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da Portaria SEI nº 002, de 25 de outubro de 2018.

Documento assinado eletronicamente por **CARLOS EDUARDO SANTOS PIRES**, **PROFESSOR 3 GRAU**, em 28/01/2021, às 16:16, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da Portaria SEI nº 002, de 25 de outubro de 2018.

Documento assinado eletronicamente por **CLAUDIO ELIZIO CALAZANS CAMPELO**, **PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 28/01/2021, às 16:17, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da Portaria SEI nº 002, de 25 de outubro de 2018.

A autenticidade deste documento pode ser conferida no site https://sei.ufcg.edu.br/autenticidade, informando o código verificador **1235966** e o código CRC **515C2631**.

Referência: Processo nº 23096.001517/2021-32 SEI nº 1235966

Universidade Federal de Campina Grande

Centro de Engenharia Elétrica e Informática

Coordenação de Pós-Graduação em Ciência da Computação

Locating Places Described in Natural Language Using Human-Machine Dialogues

José Lucas Silva Freitas

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande - Campus I como parte dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Área de Concentração: Ciência da Computação

Linha de Pesquisa: Processamento de Linguagem Natural e Sistemas de Informação Geográficos

Cláudio Elízio Calazans Campelo

(Orientador)

Campina Grande, Paraíba, Brasil

Resumo

Enquanto nos comunicamos, diversas situações que requerem que descrevamos a localização de algum ponto de interesse podem surgir. A existência de modelos que sejam capazes de interpretar estes relatos possibilitaria o desenvolvimento de diversas aplicaçõesm, como serviços de atendimento emergenciais e formas de interação com serviços de veículos autônomos. Entretanto descrições de localização feitas em linguagem natural nem sempre estão em um formato simples. Quando as pessoas fazem essas descrições frequentemente utilizam referências a pontos de interesse, lugares próximos e como estes estão espacialmente relacionados com o ponto a ser descrito. Além disso, qualquer uso de linguagem natural introduz a possibilidade de erro devido à incerteza, ambiguidade e vagueza, características frequentes nas comunicações que utilizam as linguagens humanas. A correta interpretação das expressões em linguagem natural e das relações espaciais em conjunto com os pontos de referência, assim como a projeção das regiões descritas por estes relacionamentos, representam grandes desafios no desenvolvimento de sistemas que possam localizar determinado ponto em um ambiente urbano. Este trabalho apresenta uma modelagem conceitual que busca representar conversações que transmitem informação a respeito de localização. Capturando os conceitos cruciais do domínio citado, como pontos de referência e relações espaciais, o modelo conceitual pode guiar a construção de sistemas computacionais de conversação, servindo como um conjunto de diretrizes de desenvolvimento e alertas sobre problemas comuns que podem acometer um projetista dos referidos tipos de sistemas. Como as relações espaciais são um dos conceitos mais importantes na comunicação espacial, um estudo que buscou entender seus usos pelas pessoas na linguagem do dia a dia foi conduzido. Como resultado desta investigação, um conjunto de algoritmos para projetar as relações espaciais mais utilizadas é proposto. Estes procedimentos recebem como entrada um identificador de um ponto de referência e produzem um polígono que representa a região descrita pela relação em questão. Por fim, um estudo de caso é apresentado, onde o modelo conceitual proposto é utilizado para o desenvolvimento de um *chatbot*. Tendo como escopo a cidade de Campina Grande na Paraíba, este *chatbot* faz uso dos algoritmos de relações espaciais citados para tentar localizar através do diálogo, pontos sendo descritos em um contexto urbano.

Abstract

While humans communicate, several situations that require the description of the location of some landmark may arise. For instance, when reporting some occurrence to the authorities over the phone. The existence of models that are capable of interpreting these descriptions, locating an object in space, has the potential to allow the development of many applications that make use of this type of information such as robots that can assist emergency dispatch services, or even new interfaces for existing products, such as services based on autonomous vehicles and web mapping applications. However, natural language location descriptions are not always in a clean and simple format such as street name and number. In daily conversation, people often tend to reference points of interest, nearby landmarks and their relations to the location being described; Moreover, the usage of natural language introduces the possibility of error due to uncertainty, ambiguity and vagueness, natural aspects of communication that makes use of the human natural languages. The correct interpretation of natural language expressions, of spatial relations and landmarks, as well as the projection of the regions described by these relationships, represent great challenges to the development of geographic aware systems. This work presents a conceptual model that seeks to represent conversations that convey information about location. Capturing the crucial concepts of the mentioned domain, such as landmarks and spatial relations, the conceptual model may guide the construction of computational dialogue systems by working as a set of development guidelines and alerting for common pitfalls that can befall system designers. Being spatial relations some of the most important concepts in spatial communication, a study that sought to better comprehend the usage of such relations in language by people was conducted. As a result of this investigation a set of algorithms that project the spatial extents described by the spatial relations that are most often used by people is proposed. These procedures take as input an identifier for a landmark used as reference and produce a polygon that represents the region described by the relation. Finally, a case study is presented where the conceptual model was used to support the development of a chatbot. Having as scope the city of Campina Grande at Paraíba, this chatbot makes use of the proposed algorithms to try and locate through dialogue, points being described in an urban scenario.

Agradecimentos

Agradeço ao professor e orientador Cláudio Campelo pelos ensinamentos, paciência e confiança. Agradeço imensamente por todo o essencial apoio durante estes últimos anos.

Agradeço aos professores, membros da banca, pela disponibilidade e contribuição com o meu trabalho.

Agradeço à CAPES pelo suporte financeiro necessário para a realização das atividades e aos funcionários da COPIN, pela duradoura prestatividade.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

Languages in spoken and written form are some of the most primitive and at the same time complex, communication methods that humans ever came up with. We use them to communicate our findings, wishes, to create relationships with each other and even to do politics. Natural languages are the tools that allow creation, transmission and storage of the human culture and experience.

In daily communication, we often face ourselves with a situation in which the need to reference certain places in the physical space is present. While talking about some event, contextualization through the explanation of where it happened may be necessary. The automatic interpretation of location descriptions is still an open and relevant task in the study of natural language processing and geographical information systems and accomplishing this goal would certainly enable the countless advances for many services. However, any person that has already been through this type of conversation knows that describing locations is not always an easy task. In different scenarios and contexts, people generate different descriptions about their locations [38]. Several studies already tackled such diversity in location descriptions [21, 38, 54]. The experiments made by Chagqing Zhou et. al. [54] demonstrate that people take many elements into account when describing locations, such as communication purpose, if the other person is known to them or knows the region and the level of privacy in the situation. Besides all these complexities, people still can express themselves in a vague, ambiguous, or uncertain manner.

Three problems that can happen in communication and have been explored in the literature are vagueness, uncertainty and ambiguity. These are issues that are constantly present in dialogue between people because of the nature of the natural languages. For this reason, this work explores their impact on location descriptions. Due to a reduced scope and lack of time, other language related problems such as generality were not explored but could be featured in future experiments concerning spatial language.

Vagueness can be seen as an inherent property to natural languages. It manifests itself in the synthesis of sentences that fail to convey the meaning of ideas for some of its terms may present a continuous space of interpretations. The phrase “Follow a long distance through this street”, for instance, is vague because there is a wide range of possible distances that can be classified as “long”. With no access to additional information, communication is hindered and the original idea can be distorted. As another example, in the sentence “Right next to the supermarket”, even though we might know the supermarket that is being referenced, the expression “right next to” can be applied to more than one of its sides. It is important to notice that vagueness is different from uncertainty for it is not a problem that arises from a lack of understanding about the world, but from the lack of a threshold to specify the applicability of some linguistic expressions. Vagueness is pervasive in the context of spatial information, where many terms possess an array of possible interpretations, such as “near” and “far” [2].

According to In Merriam-Webster’s collegiate dictionary[1] an ambiguous term can be defined as a word or expression that can be understood in two or more possible ways. Despite also being consequence of vague understanding of terms, in ambiguity sentences and expressions are enough to transfer the information. However, they can assume a few distinct meanings. As an example, the words “Austin” and *“Registro”*. The former can be a popular name for boys in the United States or the capital city of state of Texas, while the latter (in Portuguese) may refer to a Portuguese word related to documents or the name of a city in the state of *São Paulo*.

1.1 Motivation

In our daily lives, when we lack information to understand a message, the solution is often the simpler one, questioning. A well formed question might provide the additional information

[1] https://www.merriam-webster.com/dictionary/ambiguity

needed to solve communication problems. In fact, dialogues have been explored as tools to aid navigation systems [52].

In 1950 the English mathematician, logician, cryptanalyst and computer scientist Alan Turing, published a paper [46] in which the following question is addressed: “Can machines think?” He states that any rationalization on this matter must necessarily begin by searching the definitions of the words “machine” and “think”. Due to the abstraction level of these terms, the question can be expressed in the form of a game most known as “The imitation game”. It evaluates the intelligence of computational systems through their ability to sustain a dialogue with a human. This challenge fostered the development of several dialogue systems, the *chatterbots* or simply *chatbots* [50, 51]. These computational dialogue agents have already been successfully employed in dealing with problems in many areas such as aquatic disaster management [45], language learning [23] and even psychiatric counseling [33].

The existence of conversational agents that can interpret spatial information and in particular location and route descriptions, could represent a watershed for the development of future intelligent systems, such as autonomous vehicles and emergency services.

Current solutions for taxi services based on autonomous vehicles rely on the use of smartphones by the passengers, who must specify a precise location on classic map- or text-based interfaces. However, passengers may not have their smartphones on hand at the time they need a taxi, or may prefer not to use this kind of interface, if another efficient conversational-based and hands-free interface is available. The latter may be the case, especially, for elderly or impaired people, for example.

Emergency services may also be dramatically benefited from the availability of mechanisms capable of efficiently interpreting location descriptions. It is known that police, fire and emergency medical dispatchers process a large number of incident reports daily, normally containing location descriptions. Moreover, they must be able to maintain a conversation with a caller at the same time they are typing information into a computer [34]. However, in unforeseen situations where there is a massive surge in calls putting strain on emergency call centers, such as in big disasters, robots might be helpful in performing pre-service duties or processing recurring cases, for example.

1.2 Scope

While planning experiments a few decisions have to be made and also as a consequence of limited resources, the scope of the research is limited. An important factor is that the places that are displayed to volunteers have been carefully chosen to form a diverse set based on certain criteria. The expertise of the researchers about these places is therefore, of utmost importance for this selection of places. For these reasons, the geographic data used in the experiments as well as the datasets produced in them, both correspond to the city of Campina Grande - Paraíba in Brazil. This is a medium sized city in the northeast region of the country, with a population of over 400 thousand inhabitants. It is also considered as one of the main industrial, technological and educational centers of the region. As a consequence of this decision, all the experiments, their instructions, tools involved, spatial terms related and all the data produced were written in Brazilian Portuguese. As a final note, some of the experiments require that participants are present in the same room with the researchers, so that it is clear that no alternative sources of information (e.g. Google Maps) have been consulted, therefore most of the participants dwell in the city or other close places in the region.

1.3 Research Objectives

In an effort to foster the development of systems that are able to address the problem of locating objects in an urban space through conversation with humans, the main objective of this research is to provide two conceptual models on the domain of spatial relationships used in daily discourse and natural language dialogue aimed at conveying information about the location of arbitrary objects in urban scenarios.

1.4 Specific Objectives

The following items translate the specific goals that had to be achieved in the process of developing the research.

- Analysis of a location description corpus, in order to better comprehend the way people think and talk about space.

- Conceptual modeling of spatial dialogue between people.
- Implementation of a dialogue system prototype as a case study of one of the proposed conceptual models of the problem domain.

1.5 Contributions

Other than the aforementioned research objectives, while pursuing the research goal, the following performed tasks represent other specific contributions made during the execution of this work:

- Exploratory data analysis designed to better comprehend the way people reason about places, including quantitative and qualitative aspects of spatial relations, references to landmarks and the impact that different variables have in descriptions of location;
- Conduction of an experiment that tries to capture the way people think about spatial relations mentioned in a dialogue scenario;
- Creation of a dataset containing polygons depicting the representation of spatial relations in the minds of experiment participants. It was made publicly available, hoping to enable future work in the field;
- Proposal of a set of algorithms that implement what seems to be, according to previous experiments, the group of most frequently used spatial relations when people are faced with the job of describing the location of an object in an urban scenario;
- Proposal of a road-map with important concepts to assist the development of conversational systems that can be applied to the task of interpreting location descriptions and fine tuning this interpretation by making use of dialogue;

1.5.1 Bibliographic Contributions

During the course of the research, two papers were produced:

- "Human Spatial Reasoning in Everyday Language: Inferring Regions that Describe Spatial Relations". Published in the XXI Brazilian Symposium on Geoinformatics (GEOINFO) [18].

- "How do People Describe Places? An Exploratory Analysis of Location Descriptions in Urban Scenarios". Submitted to the Spatial Cognition and Computation Journal[2] and currently being reviewed.

1.6 Document Structure

Chapters in this document are presented in the following order: Chapter 2 lists and explores crucial concepts in the context of this research. Chapter 3 explores the relevant literature, in search of research efforts in related areas. Chapter 4 goes over an exploratory data analysis, that was possible due to an experiment that collected real location descriptions provided by people. Many important aspects of these descriptions are then analyzed, generating important conclusions about the way humans think about space and use language to express this reasoning. In Chapter 5, a set of algorithms is proposed. They are capable of generating geometries representing the spatial extents described by an important group of spatial relations and associated landmarks. Their accuracy is then evaluated through different performance metrics. Chapter 6 proposes conceptual models that try to capture the nature of conversations in which someone tries to convey the location of an object in an urban scenario. These models can be used to enhance understanding of the system being represented and aid the development of computational solutions based on the domain, such as conversational agents. Finally, Chapter 7 describes the experience of developing a conversational agent that deals with the concepts in one of the proposed conceptual models.

[2] https://www.tandfonline.com/toc/hscc20/current

Chapter 2

Background

In this chapter, the most relevant concepts to this study are explained. Section 2.1 goes over what are conceptual models. Section 2.2 describe chatbots and explores their history. Section 2.3 delves into one of the most important components of spatial language, spatial relations. Finally, Section 2.4 tackles some of the main problems one might face, when dealing with natural languages, Vagueness, Uncertainty and Ambiguity.

2.1 Conceptual Models

Conceptual Models can be understood as mental representations of physical events on the world that can be used for understanding, communication and problem solving. Concepts are the atoms of conceptual mental representations and are formed through experience, capturing observable properties of the studied events.

In the field of software development, a conceptual model can be used to represent the entities that participate in an event and the existing relationships among them. They often assist in the development and documentation of systems and database schemas. As an example, one could consider in a school system domain, the relationship between the entities "Student", "Class" and "Professor". Such models can be expressed by written text or some convention of visual representation.

It is important to notice that it is impossible to model the real world, therefore conceptual models try to represent our conceptualization of the world. As a consequence of this fact, the evaluation of such models is a difficult task. Their requirements can be expressed as completeness (coverage of involved concepts) and soundness of the concepts.

2.1.1 An Observation on Notation

The conceptual models depicted in this thesis, represent all the concepts using orange ellipsis. The continuous arrows with red boxes represent relationships between the concepts. These relationships can represent events or possession relations. In the former case, a concept acts upon another, as in a *person* checking if an *acceptance region* is small enough. In the latter case, a concept acts as attribute of another, as in a *description interpretation* having a *location description*. The dashed arrows represent an inheritance structure between concepts that denotes a relation of textitis a type of. For instance, a *further inquiry* is a type of *message*. Green boxes depict the cardinality of the relations. The cardinalities in the *sender* relationship can be read as: “A message has at least one and a maximum of one sender, while a person can be a sender in a minimum of zero, and a maximum of n messages”.

An example of a simple conceptual model that represents a university course, with students and professors is presented in Figure 1. In this model the concepts are **Person**, **Student**, **Professor** and **Course**. A person, can be of sub-type student or professor. A student takes a course, and the cardinality of this relation suggest that one instance of the student class can take many courses, while a single course can be taken by many students. A professor teaches one course, and the cardinality implies that a professor can teach many courses while a course can only be taught by a single professor.

2.2 Chatbots

Humans communicate through the usage of natural language. New interfaces to computer systems, particularly the ones that make use of natural language, have the potential to enable ease-of-use for users in performing complex human-computer interactions [53].

Chatbots are conversational systems that can interact to users through natural language, mimicking communication between humans. The first chatbot was developed in 1996, used pattern matching and substitution to simulate dialogue [51] and performed surprisingly well. Since then, significant advancements in the area have been made, especially with the recent developments in the fields of Natural Language Processing, Natural Language Generation

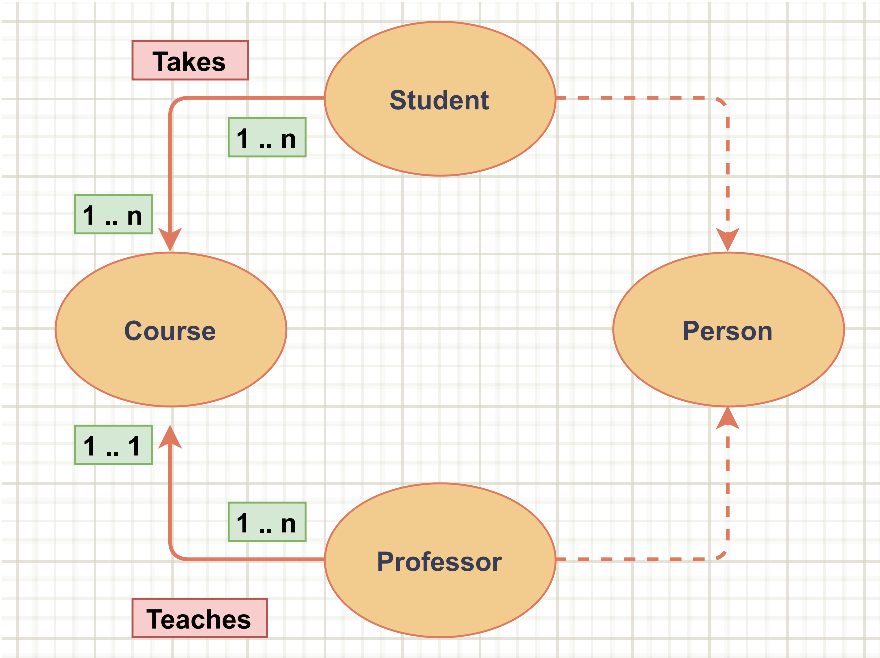


Figure 1: Example of a conceptual model that represents students, professors and courses.

and Machine Learning in general.

These conversational systems can be complex for their development includes many non-trivial tasks such as intent identification, information extraction, dialogue management and response generation.

2.2.1 Natural Language Processing

In any conversational agent, the first step in the dialogue process is to process incoming messages in search for relevant information. In this stage, messages are usually in the unstructured form of natural language and therefore need to be processed in order to generate data structures containing the semantic information that is needed in the dialogue management phase. The type of semantic information that is extracted from the text depends heavily on the selected type of dialogue management.

Keyword Detection

Some agents may react in a determinate way given the presence of a given keyword in the input text. In fact, these key expressions can be more than simple words. An agent can define sets of combinations with an arbitrary number of words or even patterns of expressions (using strategies like regular expressions) to which it responds to. For this reason, keyword detection can be a time consuming task. Old chatbots like the aforementioned Eliza rely heavily on keyword detection strategies.

Intent Identification

As systems become more and more sophisticated, the need for domain specific agents rise. In this type of agent, the goal of the message provided by the human becomes crucial in the dialogue management. Instead of only knowing that a utterance is a question, it becomes necessary to understand if the user wants to order a pizza or only check the prices. This purpose of the message in the specific context that the chatbot is supposed to work is usually called “intent”.

The intent identification process can be done by searching for keywords but this is often not an accurate way of approaching the problem. With the availability of conversation data, machine learning models can be employed. The relevant intents for the domain assume the role of the classes that the classifier assigns a message to.

Statistical models that have been used to perform intent classification are Support Vector Machines (SVM) [12], Conditional Random Fields (CRF) [26] and various Deep Learning neural network architectures. The training can be done with many features from the messages, such as bag of words (i.e. number of occurrences of each word), part of speech tagging (i.e. each word is labeled with a grammatical class such as nouns, adjectives and etc) and the presence of relevant named entities.

Entity Extraction

When a person orders a pizza, after detecting the message intent, an important next step is to find out which toppings and sauce. Traditional rule-based assistants tend to explicitly ask for this type of information in turns, one at a time. Even though this approach works, it does not resemble a natural human dialogue flow since the information can already be available in the original message. For this reason, the extraction of domain specific entities needs to be carried out for chatbots that are capable of having more natural conversations.

Machine learning models such as the ones mentioned in the previous section do a good job in extracting relevant entities from textual data. In fact there are already many models available on the internet, that are capable of recognizing common entities such as places, people names and dates.

Despite the aforementioned advancements, mimicking human dialogue abilities is still a difficult project. Most of the commercial chatbots in use nowadays are restricted to closed domains, where they are only able to interpret references to entities that belong to a single domain. Examples are the ones available on Facebook messenger, such as the one by Whole Foods Market [1] that lets users search for recipes and the one by Uber [2] that gives users the possibility to request a ride without using the Uber app.

2.2.2 Dialogue Management

Probably the most important task for any agent is the dialogue management. It is what makes an agent conversational. Assuming an input in the form of pre-processed textual data, the dialogue management utility selects the appropriate action that the bot should take and produces a response to be sent back to the user. More complex agents as the ones mentioned in Section 2.2.1 are able to make use of external data sources such as databases or search engines before they decide what to say.

Concerning dialogue management there are usually two types of chatbots [36]. The rule-based [49, 51] and the ones that make use of generational models [13].

Rule-Based Chatbots

In 2012, Luka Bradeško and Dunja Mladenić conducted a survey [3] to analyse the chat-bot systems awarded with the Loebner Prize[3], an annual competition that rewards the most human-like conversational agents. The authors found that most of them are still based on pattern matching and rules.

Rule-based chatbots have access to a list of rules in the form of: **Pattern → Response template**. If an input message matches a certain pattern, the bot produces the response using the appropriate template and sends it back to the user. Patterns can include keyword matching or even entire sentences.

[1] https://www.wholefoodsmarket.com/

[2] https://www.uber.com/

[3] https://en.wikipedia.org/wiki/Loebner_Prize

This type of bot is restricted to the flow of conversation that is programmed by the developer, this makes it hard to add new features and easy to fail when the user inputs an out-of-script message.

Generational Models

Instead of relying on a database of predetermined answers, generational models build all responses while the conversation takes place. This is usually done through the usage of machine learning models trained with real conversation data. These models usually belong to a class of "machine translation" techniques. In traditional machine translation problems, models are trained with a corpus containing text in more than one language and are capable of translating text between these languages.

A variation of statistical machine translation techniques is often used in the context of conversational systems. In this variation, machine translation models that can translate from an input in one language to an output in the same language are used. One example of such generative strategies are sequence to sequence models (seq2seq) [41], a machine learning strategy that make use of two recurrent neural networks to encode and decode an input sequence, producing a response.

2.3 Spatial Relations

One piece of crucial information often used by people when describing the location of an object in an environment are spatial relations. They describe how an object is located in a given scenario, in relation to another reference object. In the sentence "The place is near the school, right next to that big old church", **near** and **next to** are the parts that play the role of spatial relations, as they describe the location of the place in relation to the school and the big old church, respectively.

Spatial Relations have been studied for decades and have been classified into metric, topological and projective [4]. Metric relations define positioning in the context of a scalar quantity. They often specify how far apart are the involved objects. Examples of metric relations are "Closer", "far away" and "nearby". While metric relations are important, humans seem to have a qualitative reasoning of space [11].

Topological relations describe the position of objects in terms of the intersections of their interiors, boundaries and exteriors. They have been extensively studied [10, 16, 28], and in fact are even supported by spatial query languages. Examples of topological relations can be seen in Figure 2.

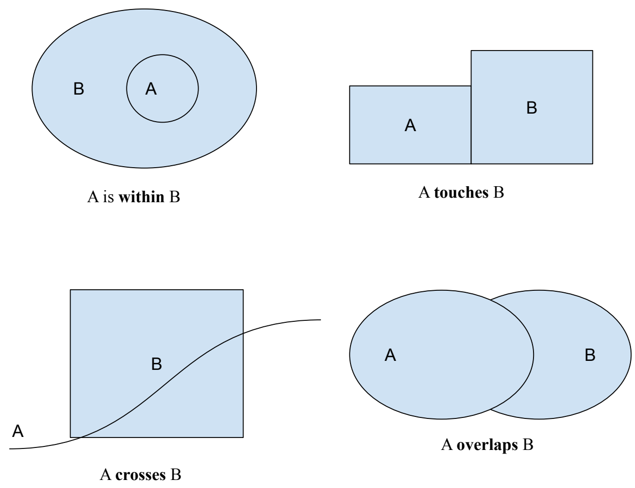


Figure 2: Examples of Topological Relations

Directional relations are a common subcategory of projective relations that describe where an object is located in relation to other objects using qualitative terms. Examples of directional relations can be seen in a sentence such as “The building to the right of the elementary school.”. Here the expression “to the right” represents the directional spatial relation. In fact, the set of directional relations, includes expressions used in natural language such as “right of”, “in front of” and “between”. Examples of directional relations can be seen in Figure 3.

2.3.1 Frames of Reference

When someone say “The car is to the right of the university”, the message can be interpreted in different ways according to one’s spatial mental reasoning. If directions are defined based on the observer’s point of view, the region defined by the relation may assume a completely opposite position than if directions were defined by the position of the reference object itself.

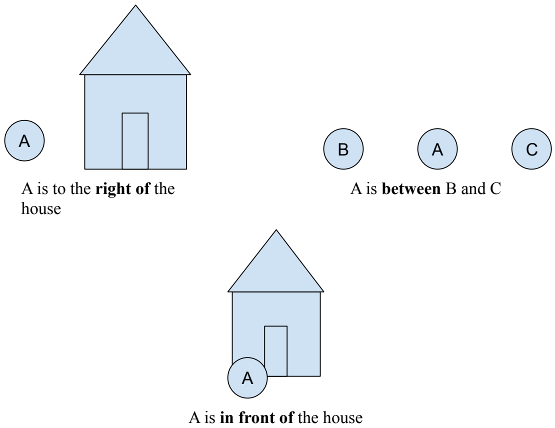


Figure 3: Examples of Directional Relations

Contextual information such as a clear frame of reference is then needed to disambiguate the sentence when a few relations are used. Frames of reference are then theoretical constructs that help disambiguate spatial relations and have been classified into **intrinsic** where an intrinsic property of the reference object (such as its front) defines orientation, **extrinsic** where orientation is defined by another external landmark and **deictic** that defines orientation based on an observer's point of view [37]. Figure 4 depicts an example of a relation being defined by a deictic frame of reference.

2.4 Vagueness, Uncertainty and Ambiguity

Certain linguistic expressions carry the potential to hinder communication due to the phenomena of vagueness, uncertainty and ambiguity.

The predicate "child" certainly can be applied to a person who is four years old, while the same cannot be said to someone in their eighties. However, the applicability of the said predicate is uncertain when considering a person who is 12 years and 11 months old. This confusion is due to the vague nature of the "child" predicate. A term can be said to be vague, when its applicability lacks a clear boundary case (i.e. at which age, a person can no longer

Figure 4: The eye represents an observer positioned in front of the landmark (car dealership). The orange polygons are thus positioned to the right of the landmark, from the observer point of view.

be considered as being a child?).

Vagueness can be expressed through an ancient puzzle called the *Sorites Paradox* [4]. Credited to the Greek philosopher Eubulides, the puzzle describes a paradox that arises from the lack of a clear line that divides the applicability of a vague term such as "heap". One grain certainly does not constitute a heap, the same way two grains also do not; if two grains do not form a heap, it follows that three also do not and so on. This leads to the absurd conclusion that no number of grains can ever form a heap. The problem can be formally expressed using the Conditional Sorites, expressed in the conditional form 2.1.

[4]https://plato.stanford.edu/entries/sorites-paradox/

$$
\begin{array}{r}
NH_1 \\
NH_1 \rightarrow NH_2 \\
NH_2 \rightarrow NH_3 \\
\cdots \\
NH_{n-1} \rightarrow NH_n \\
\hline
NH_n
\end{array}
\tag{2.1}
$$

Let NH_n be a predicate that means "n grains are not a heap". Using *modus ponens*, it seems that the deduction that no amount of grains can be considered as a heap is valid.

Vagueness is pervasive in spatial language, in fact most metric spatial relationships can be considered vague. Suppose someone says that a point A is "close to" another point B. Even in possession of an accurate measurement of the distance between the two points, it is possible that the applicability of the spatial relation may yet be uncertain. This happens because of the vague nature of the expression "close to". Other examples of vague terms are "near" and "far away/distant".

Brandon Bennett [2] classifies a proposition as uncertain, when the reason we do not know whether it is true or not arises from the fact that we do not have complete and accurate knowledge about the world. For example, one cannot simply state if it will rain tomorrow, therefore we are uncertain about this topic. Uncertainty is then distinguished from vagueness as the former do not rise from a lack of understanding about the world. In the spatial context, uncertainty can be observed when the speaker references unknown places.

Ambiguity refers to the property of terms that given a certain context, allow more than one possible interpretation. Thus, ambiguity generates uncertainty, as an ambiguous statement cannot easily be verified to be true or not. The word "court" can take the meaning of a place where justice is administered or that of an open area (i.e. a food court, or a game court).

The distinction between ambiguity and vagueness is often debated. In ambiguous expressions different interpretations are possible depending on contextual information while with vague predicates, forming any interpretation is complicated due to the lack of a clear boundary case of applicability.

The next chapter will go over the relevant work in the literature that is related to the many stages of this research such as the exploration of the spatial reasoning by humans, proposal of algorithms to project spatial relations and conceptual modeling in the communication and spatial information scenarios.

Chapter 3

Related Work

The idea of modeling the real world phenomena of conversations related to the location of places in an urban space is to the best of our knowledge, a new endeavor. Such is the case for the application of such models. The field of conversational agents that can make use of spatial information is still uncharted territory. In fact, Frei [17] has demonstrated that it is possible to integrate a conversational agent, to a geographical information system in a meaningful manner. It is expected that this work will assist new experiences in this scenario, by providing researchers and developers with a model that tries to capture the way people reason about the task of describing the location of an object in space.

Even though not many studies have dealt with the particular problem that this thesis tries to address, some of its contributions have been previously explored. Such is the case for the experiments that try to shed some light into the way people think about space and spatial relations, which have been the target of many studies in the last few decades. Conceptual Models of events related to communication and spatial information have also been explored in the literature.

3.1 Human Reasoning About Space

In the area of Geographic Information Retrieval (GIR), many efforts have been made towards methods of automatic extraction of geospatial information from textual data [24,25,40]. Proposed approaches include the identification of terms that express references to place names and to physical entities. On the other hand, the work done in this thesis is concerned with human descriptions that convey where these entities are located in an urban setting.

Human reasoning about place and related spatial concepts is a topic that has been previously explored [20, 28, 38, 54]. There is some intersection between the aspects of place descriptions that other research have explored and those that we are interested in, however, the current study differs in aim, scope or methodology. This work combines many analysis of different aspects of location descriptions, such as their types, usage of spatial relations and landmarks; and the effect of variables like age, gender and knowledge of the area.

Vasardani and Winter [48] raised the question of whether places can be identified through a set of properties encoding the concept. They suggested Alexander's 15 structural properties [1] and examined how they correspond to properties of different applications of place. In a subsequent work [47], they conducted a case study to explore whether these properties are cognitively supported, aiming that a subset (or a superset) of them could be adopted to define computational representations of place instances, based on a place constructor specified as a function of place properties (i.e., attributes). The authors analysed place descriptions (16 urban, 13 rural and 13 indoor place descriptions) provided by 14 graduate students of the University of Melbourne.

Although the study of Vasardani and Winter [47, 48] is noticeably related to the exploration presented in this thesis, it also differs in several aspects. While the authors asked the participants to imagine any existing places (of their choice) and to describe them in general terms, we defined the places the participants had to describe, aiming at analyzing their behaviors in different situations (e.g., known vs unknown places). Moreover, the participants of the experiments described here had to provide descriptions based on their interactions with a virtual environment, which simulates their views of the environment. In fact, the concept of place in this work does not refer to those that can be associated with named places, such as parks or squares. Differently, we consider punctual places in a city environment, that can be described in terms of existing places in their surroundings. For this reason, we prefer to refer to them as *locations*. Clearly, the aims are different: while they aim at defining place constructors based on a set of properties, this study is aimed at examining how different people describe the same set of selected places/locations, with the hope of getting meaningful insights about features (e.g., age, gender) that may be related to different behaviors.

Zhou et al. [54] investigated the types of descriptions produced by people. The focus is on identifying if descriptions are tailored to different audiences, and the factors that influence these adaptations. The analysis was done through interviews, where participants had to answer how they would describe their location to different people, and what would make them change the descriptions. The locations were selected among those the participant have reported to have recently visited. The results suggest a few types of descriptions, such as *generic*, *well-known public*, *specific public*, *personal* and *activity based*. Factors influencing descriptions were also found, among them are the purpose of the message, if the recipient knows the sender of the message and if the recipient knows the area. The work of Zhou et al. [54] can be considered similar to some of the analysis described in this thesis in the sense that both aim at exploring the effect of certain factors in location descriptions. However, the studies are considerably different in terms of more specific objectives and in relation to methodological aspects. In particular, in the current research, the locations were carefully selected based on their characteristics, and all participants had to provide descriptions for them, allowing us to compare the results. This allowed the study of the variation of descriptions provided by the participants, to the exact same locations, as opposed to each participant describing a different place.

Richter et al. [38] also investigated types of place descriptions. A corpus of place descriptions collected through a mobile game was explored through a clustering algorithm to group descriptions based on a few features such as granularity of the elements, presence of indoor/outdoor references, use of spatial relations and description style. Assuming that only a few number of types exists, the results show three prevalent classes of descriptions: location, locomotion and route descriptions.

Tomko and Winter [44] also investigated types of place descriptions and proposed two concepts that are similar to the ones found by Richter et al. [38]: destination and route descriptions. Respectively, these types of descriptions are concerned with “where" places are and “how" to get there. An example of each class can be read in Table 3.1. Destination descriptions locate places based on reference points in the vicinity. Route descriptions are composed of step-by-step instructions to reach the place starting from an initial point. While some descriptions can still combine characteristics from both categories, it is postulated that destination descriptions are usually shorter, therefore the cognitive workload of producing them is smaller. Considering this classification, the present study tries to identify which

Destination	The place is near the lake, in front of the church.
Route	Follow the road and turn left after the church, the place is at the right side of the street.

Table 3.1: Examples of Destination and Route Descriptions

category is used more often by people.

Finally, making use of MS-MARCO [32], a machine reading comprehension dataset provided by Microsoft, Hamzei, Winter and Tomko [20] conducted an analysis of the relationship between questions and human-generated answers by type, scale and prominence of places referred to. Their results indicate that the answers generated by humans follow a specific pattern and that the type, scale and prominence of places have a direct relation to the answers. The authors analysed data containing geographic references at different levels of granularity, and observed that a few types of references (e.g., states) are more frequently referred to. In this thesis, the scope is limited to references in a sub-city level.

It is also worth mentioning that related research has been conducted in the field of robotics, especially in topics related to the automatic interpretation of human descriptions of place and location [22, 29, 43]. The descriptions are particularly important for the development of robots that are capable of understanding human communication in the context of navigation through an environment. On the other hand, the present study contrasts with the robotics literature in the sense that its focus is on the ways people naturally produce descriptions, in a human-human setting. Moreover, existing approaches in the field are normally concerned with indoor or other controlled environments (e.g., [14, 42]), rather than a more complex environment like a city.

However, the findings of this study may contribute to the development of more robust human-machine algorithms, capable of choosing the appropriate language resources according to the situation. For instance, an autonomous vehicle may opt for different language resources while communicating with people of different ages.

3.2 Spatial Relations

Previous research has investigated the contrast in the way people reason about spatial relations expressed through visual or linguistic resources. Mark and Egenhofer [28] have conducted a study aiming at comprehending how people reason about spatial relations involving lines and regions. They tried to observe if people would group together diverse drawings of lines and regions representing spatial relations between roads and parks respectively. Participants were presented with sentences describing groups of drawings, and were asked to rate how much they agreed that the sentence actually describe the relation in the sketches. The study concluded that most of the relationships identified by people are among the 19 described by the 9-Intersection model of topological relations [16].

While their work presents meaningful insights into how people perceive some spatial relations, it is worth verifying how these relations are observed in a more complex environment such as a city, where many other geometric features are present, and considering not exclusively line-region relations. Furthermore, rather than rating the agreement between relations expressed textually and visually, in this study the focus is on exploring descriptions provided by the participants themselves. In this way, it is expect to get a grasp of general preferences about spatial relations adopted to formulate location descriptions in an urban scenario.

Considering the types of spatial relations, while metric relations are important, humans seem to have a qualitative reasoning of space [11]. Despite the fact that topological relations have been extensively explored in the literature, this specific type of spatial relation does not seem to be what comes to the minds of people when trying to describe the location of objects in an urban space. In fact, it might be hard for a person to accurately define clear lines defining the interior, boundaries and exteriors of some landmarks that can be found in many cities. For this reason, this thesis is focused in finding out the spatial relations that are present in location descriptions provided by people in a real location description scenario. For this reason, although not all of them, most of the relations explored in this work fall in the directional category.

Directional relations such as “right of”, “in front of” and “between” are ambiguous and need additional contextual information such as Frames of Reference [7] to be accurately interpreted. In his work, Clementini defines a taxonomy of frames of reference, mapping relations to the 5-Intersection model of projective relations [9], this gives the additional geometric definitions needed to compute relations. [8] build on top of this mapping and present a Java application framework that implements the directional relations given the assumption that the relations are being interpreted in a few of the frames of reference.

In the present work, a different approach is presented in computing directional relations. A set of algorithms that generate regions that correspond to the relations by computing intersections between buffers around landmarks and nearby streets is proposed. The idea is that these procedures could be used in an application after a stage of entity extraction from natural language, where landmarks and spatial relations are collected, to generate possible projections of spatial relations. As an example, a conversational agent could extract references to landmarks and spatial relations from a textual place description and project the region being described, applying the appropriate algorithm to the landmark being referenced. Despite most of the relations explored being directional, the set of relations covered by the proposed algorithms is not intended to be a exhaustive list of all relations in this class. In fact, the main focus of the study is to explore a subset of relations, among the ones that, according to our findings, are most often used when people describe places. This subset includes common expressions that although are widely used by people in conversation, to the best of our knowledge have neither been categorized as directional nor explored before such as Next-To, Near and At-Street.

3.3 Conceptual Models of Space and Communication

In 1989 Mark and David M. [27] explored the cognitive process behind decisions during vehicle navigation and driving in large-scale geographic spaces. Their work delineate terms and concepts concerning vehicle navigation and proposes a conceptual model whose purpose is to assist in designing components of navigation systems.

The work by Mark and David M. presents relevant insights into the cognitive process in driving situations. As a result of the focus on navigation systems, their model is concerned with procedures such as route planning and generation of instructions. This thesis is focused in the exploration of location descriptions that a human produces to another. In a scenario like this, humans make use of the knowledge that the other person may possess about the particular space to reduce the cognitive load of producing step-by-step directions to reach a location. Route planning and navigation can thus be employed as a posterior process that takes place after a destination is resolved.

Distances and metric spatial information were the main concern of the work by Montello and Daniel R. [30]. The authors review the literature and propose a conceptual model that tries to explain human cognitive processing of distance information. The work objective is to increase the understanding of how humans think about distance. A classification of processes and information sources is then proposed. It states that people process information about distances in the environment through one or more of four classes of processes: Working Memory, Non-Mediated, Hybrid and Simple-Retrieval. This processing also involves three sources of information: Number of Environmental Features, Travel Time and Travel Effort.

Even though the current thesis shares with the work by Montello and Daniel R. the goal of improving the understanding of how humans think about space, the studies differ in the object of exploration. Their work is focused on quantitative aspects of spatial reasoning in the form of computing distances, while this thesis is focused on a qualitative fact of the human reasoning about the physical space.

The study by Gryl et al. [19] is also focused on qualitative aspects of spatial reasoning. The authors propose a conceptual framework for dealing with spatial information. Their work then presents a categorization of the verbal expressions in route descriptions and through the analysis of a corpus of natural language descriptions, introduces concepts that are used to model the semantic content of these expressions, such as the area of influence of an object in space and the idea of displacement. One of the main goals of the study is the development of a knowledge-based system which manipulates spatial and temporal knowledge, simulating the actions that people take people while making a route description.

The conceptual representation proposed by Gryl et al. includes similar concepts to some of the ones presented in this thesis. In their work, route descriptions have two main components, landmarks and actions, the former assuming a similar meaning to the reference objects that in this work, are also named landmarks. However, both works are concerned with different types of natural language descriptions. As already mentioned in Section 3.1 place descriptions have been categorized into route descriptions and destination descriptions. Although the analysis and concepts presented in this thesis are not restricted to one particular type of place description, the corpus analyzed in this study is mostly composed of destination descriptions and thus its main concern is in this type of description as opposed to the work of Gryl et al. which is mainly focused on routes. This work is also different in the sense that it is aimed at allowing the development of dialogue systems that are capable of locating objects in space through conversation with humans.

The marketing literature is concerned with the issue of knowledge sharing, this has led to the development of conceptual models that deal with this phenomena through different mediums. Schlegelmilch et al. [39] proposed a conceptual model to represent the process of knowledge transfer in geographically disperse marketing functions of multinational companies, focused on units located in places with different cultures. The works by Panahi et al. and Cheung et al. [6, 35] try to model the sharing of knowledge through the usage of online social platforms (e.g. Twitter and Facebook).

The object of study of this thesis is also a form of knowledge transfer. In the explored dialogues, a person tries to share a knowledge about space. However, the medium through which this sharing happens in the present work is conventional conversations. Even though a great deal of attention have already been given to knowledge transfer in general, there is still a lack of studies about conceptual representations of dialogue and in particular spatial dialogue.

Table 3.2 condenses the information of this section by classifying the mentioned studies into 4 dimensions: those which aim at exploring the spatial reasoning of humans; those whose focus is on studying spatial relations; those that propose algorithms for projecting spatial relations using geographic data and those that include conceptual models of dialogue. An important thing to notice is that these works are usually separated by areas of research, this makes it rare to find studies that share all the goals of this thesis.

The first six entries in the table [20, 28, 38, 44, 48, 54] represent studies whose purpose is to delve into the way people think about space, with the last one also analyzing spatial relation use. On this dimension, Clementini [7] laid an important foundation on the study of spatial relations, with a subsequent work even proposing ways to project such relations using spatial data [8]. As previously mentioned, the set of relations for which algorithms are proposed in this thesis is different, as well as the approach into designing them. Even though some studies are concerned with spatial reasoning and also propose conceptual models [19, 27, 30], the models they propose are not related to dialogue. Finally, some studies in the fields of marketing and administration deal with knowledge sharing through several mediums. Although none of these works study verbal dialogue between people, they are the closest that could be found in the literature. Nevertheless, none of them are related to spatial information [6, 35, 39].

Study	Spatial Reasoning	Spatial Relations	Algorithms	Knowledge Sharing Models	Dialogue Models
Vasardani and Winter [48]	X				
Zhou et al. [54]	X			*	
Richter et al. [38]	X				
Tomko and Winter [44]	X				
Hamzei, Winter and Tomko [20]	X				
Mark and Egenhofer [28]	X	X			
Clementini [7]		X			
Clementini and Bellizzi [8]		X	X		
Mark and David M. [27]	X				
Montello and Daniel R. [30]	X				
Gryl et al. [19]	X				
Schlegelmilch et al. [39]				X	
Panahi et al. [35]				X	
Cheung et al. [6]				X	
This Research	X	X	X	X	X

Table 3.2: Comparison Between this thesis and the related works

The next chapter begins the exploration of the spatial reasoning in the minds of humans. An exploratory data analysis done in the dataset produced by a previous experiment is described. It examines location descriptions provided by real people, exploring the usage of landmark references, the choice of landmarks to be referenced and spatial relations.

Chapter 4

Human Reasoning About Space

One of the first endeavors of the research process was to try to get a better comprehension on how people reason about space, understanding what factors affect the making of location descriptions. A first experiment was crucial to accomplish this goal.

4.1 How do People Describe Places?

Understanding how humans devise place descriptions is decisive for the implementation of preciser systems that are capable of finding places by communicating with humans. To take a step in this direction, Neto [31] conducted an experiment to collect and investigate location descriptions. As an initial step in the present research, an exploratory analysis with the dataset produced in the experiment conducted by Neto was carried out. The data seems to show that variables like gender and age have an impact on descriptions. For instance older people tend to give lengthier descriptions. In the following sections, the methodology of the experiment by Neto is presented, then, a brief discussion on some of the analysis findings are presented.

4.1.1 Materials and Methods

This subsection briefly describes the methodology of the study executed by Neto. In the mentioned study 57 volunteers, their identities concealed, described ten different locations within the city of *Campina Grande* in Brazil. Before conducting the experiments, a few descriptions for some randomly chosen places were written by the researchers and based on the time spent on this task, an estimate was made that two minutes would be an adequate time for the descriptions to be produced without putting pressure on the participants. For this reason, they were given two minutes to produce each description. Through a questionnaire, participants informed gender; age; whether they live or have ever lived in the city; and level of knowledge about the city (in a scale of 0 to 10).

The ten locations that the participants had to describe were chosen to form a diverse set with different levels of difficulty. This was done through the selection of places located in central areas of the city and/or surrounded by landmarks, and places located in more suburban areas with few landmarks nearby which in general, are harder to describe.

Respondents were asked to imagine a situation where they were standing at the given place and needed to describe the location to a friend, so that the friend could reach the location and pick them up for a ride. To prevent the situation where participants would use a more formal language, due to the fact they were writing (instead of speaking) and because they were participating in a research experiment at a university, they were instructed to describe the locations using the language they use daily with friends and family,

To perform data collection, a Web application was developed. It presents to the volunteers an embedded iframe from Google Street View[1], displaying the place to be described (where the participant is supposed to be standing on - normally a sidewalk). The tool's user interface is shown in Figure 5. It contains a timer that displays the remaining time for each description, a text area to input the description and a button to move to the next place sample. Through the interaction with Google Street View, volunteers can navigate the surroundings of the places and visualize landmarks. By default, Google Street View displays labels on the images to help the users identify street names. To approximate the simulated environment to the real world, these labels have been removed from the images shown on the application interface, along with all links to access the Google Maps service. On the other hand, all textual information physically existing in the environment was left unchanged, such as street names written on plates fixed on buildings' walls and signs on storefronts. The produced dataset is publicly available[2], with the aim of fostering further research in the area.

[1] Google Street View - `https://en.wikipedia.org/wiki/Google_Street_View`

[2] https://github.com/jslucassf/sc-location-descriptions

Figure 5: Interface of the Web application used for data collection.

4.2 Exploratory Data Analysis

The dataset produced by the experiment performed by Neto [31] and described in Section 4.1.1 was used in an exploratory data analysis that was executed as part of the work of the researchers of this thesis. Its results represent important findings about how people describe places, and can be used to support decisions during development of many Geographic Information Systems (GIS) applications. Therefore, now this chapter is focused on this analysis.

4.2.1 Landmarks in Descriptions

In order to identify the use of landmarks and spatial relations, a machine learning classifier was trained to perform Named Entity Recognition (NER), customized for the target scenario. As an example, given an input such as *"close to the old church"* the model should identify the landmark *"old church"* and the spatial relation *"close to"*. The classifier is based on Conditional Random Fields(CRF) [26], a statistical model used in natural language processing, which is able to predict sequences of labels for sequences of input samples.

The classifier detected that in 92% of the descriptions at least one landmark was referenced. Even accounting for classification errors, this result suggests that landmarks are one of the main features used by people when describing locations of places. The data also shows that the frequency in which landmarks are referenced seems to be related to the age of the research participants, with a moderate positive Pearson's correlation of 0.50 between these two variables. No correlation was found between the usage of landmarks and self assessment of city knowledge.

The Google Places API provides information about the places referenced in the descriptions, including for each place, an average user rating and the number of ratings received. There seems to exist a strong correlation between the number of people that rated places that are close to the points that the respondents described and the usage of landmarks (0.70 Pearson's coefficient), this means that when the place to be described was in a region with landmarks that have many user ratings, participants referenced more landmarks than in regions with landmarks that have lesser user ratings.

4.2.2 Spatial Relations Usage

When describing locations people will often make use of spatial relations. These can be described as expressions that define how an object is spatially related to another. Through manual inspection and the usage of the entity extraction model, the expressions that were most commonly used by the respondents are listed in Table 4.1.

The expressions shown in Table 4.1 are not intended to be a comprehensive list of all spatial relations in the Brazilian Portuguese language. However, 79.8% of the descriptions collected in the questionnaire, contain at least one of the aforementioned terms.

4.2.3 How often each spatial relation was used by the respondents?

From the list, *in front of* is the term used more often. At the second place, the relation *near*, followed by *next to*. These three are by a large difference, the preferred relations by the respondents, as can be seen in Figure 6.

4.2.4 Where does near becomes far?

As already mentioned, vague terms are the ones that do not possess a clear line that determines the extent of their applicability. *Near*, one of the most often used spatial relations is vague for it may require contextual information for its validity to be verified. The uncertainty

Portuguese	English
Frente	In front of
Próximo, Perto, Nas proximidades	Near
Ao Lado, Vizinho	Next to
Ao Redor	Around
Às margens, Colado	At the edge, touching
Entre	Between
Antes	Before
Depois	After
Atrás, Por trás	Behind
Esquerda	Left of
Direita	Right of
Sobe, Subindo	Going up
Indo	Going
Na entrada	At the entrance

Table 4.1: Spatial Relations Names and their Translations

that emerges from vague metric relations has been studied and for instance, estimates of distance between reference objects expressed by spatial terms, change according to variables such as the size of the objects themselves [5].

Distances between the goal location and the reference landmarks when people described them as being *near*, were compared in an attempt to understand what people mean when using such a vague spatial relation.

Distances between points and regions were computed using an edge-to-edge approach, since for larger regions, the increase in distance (to the central point of the region) do not necessarily mean that the places are further away from each other as can be seen in Figures 7 and 8. In this setting, the smallest distance recorded is 8.88 meters, while the largest, 581.37 meters, as represented in Figure 9.

Most cases of the relation *near* refer to places that are at most 300 meters of distance, although there are a few cases of places more than 500 meters away (Figure 9). The median distance of the point referred as 'near' was estimated using the bootstrap resampling method

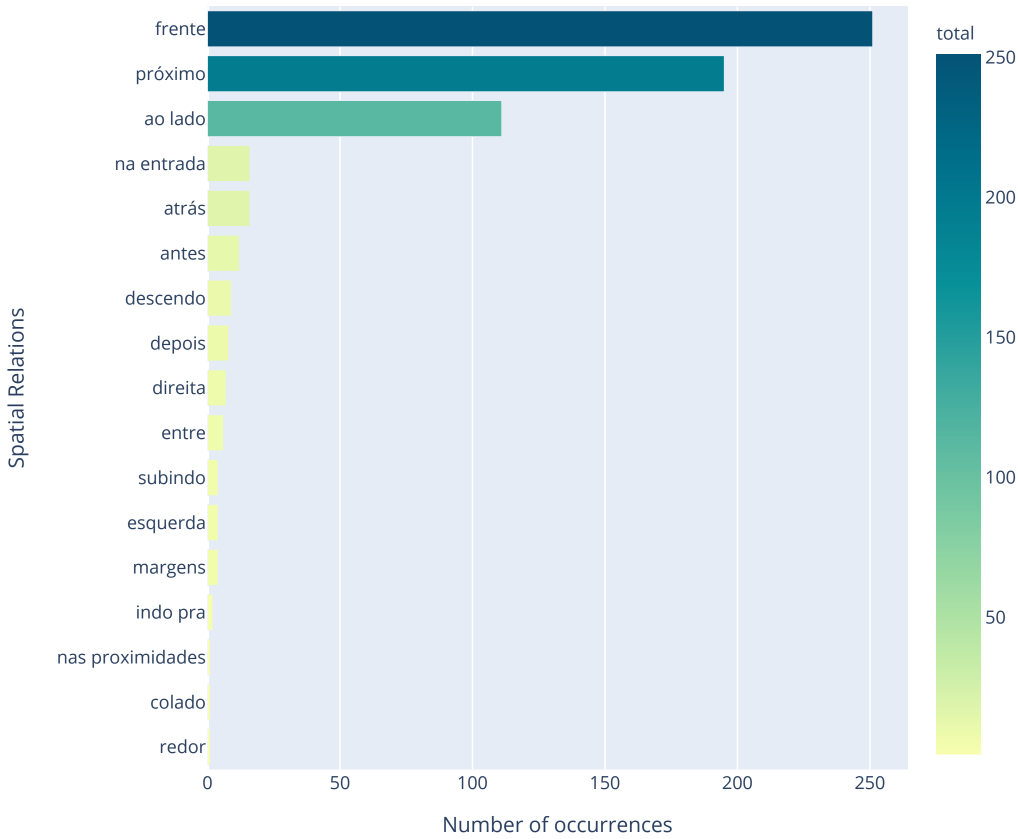


Figure 6: Frequencies of usage of spatial relations.

[15]. The data collected seems to show with a confidence level of 95%, that the median distance of a 'near' reference, lies between the range of 144 and 183 meters (95% CI [144.0, 183.0]).

4.2.5 Proximity vs Prominence

The description to two goal locations in the experiment presented some insight into the choice of landmarks. Although there were reference points that are closer to the goals, most of the time respondents preferred to choose places that are further away, but that are more prominent.

Figure 10 shows the usage of references to landmarks in the context of these two goal locations. In the first chart, the points *Supermercado do Germano* and *Fofex* are about 10

Figure 7: Edge-to-Edge distance

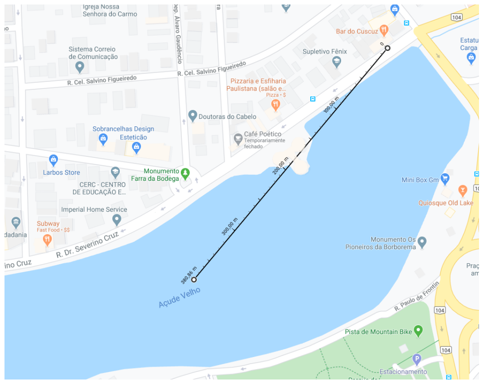


Figure 8: Edge-to-Center distance

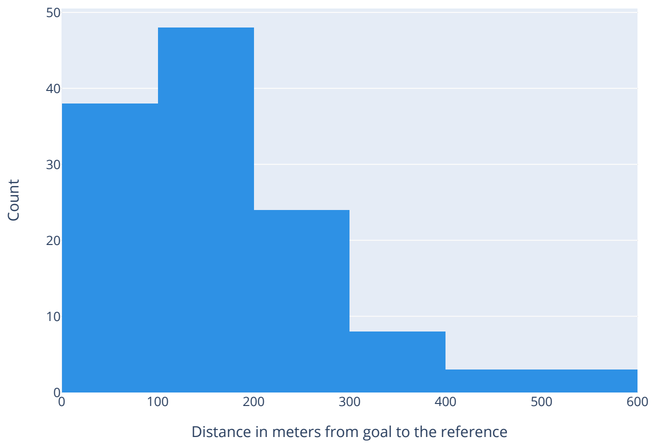
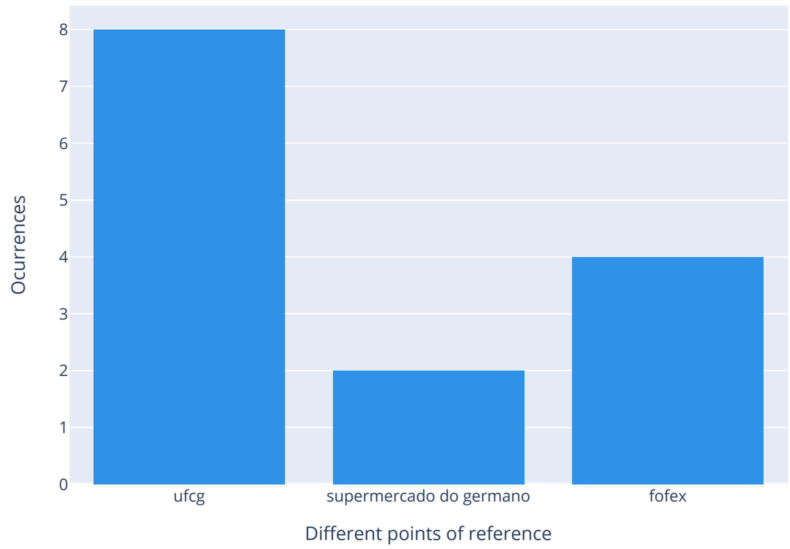


Figure 9: Frequencies of usage of distances in the context of the spatial relation near.

and 140 meters away from the goal, respectively. Nevertheless, respondents chose way more often as reference, the local Federal University which stands at more than 200 meters of distance. The city of *Campina Grande* is an important academic center in the state of Paraíba, with many universities from whom the Federal University of Campina Grande is one of the most important. For this reason, it is an important landmark known by many city residents.

The second histogram in Figure 10, displays a similar phenomena. The most frequently used landmark was *Partage Shopping*, the biggest shopping mall in the city, despite the fact that a Nissan car shop and a police station were immediately in front of the goal.

The next chapter further explores the most frequently used relations mentioned in Section 4.2.3.



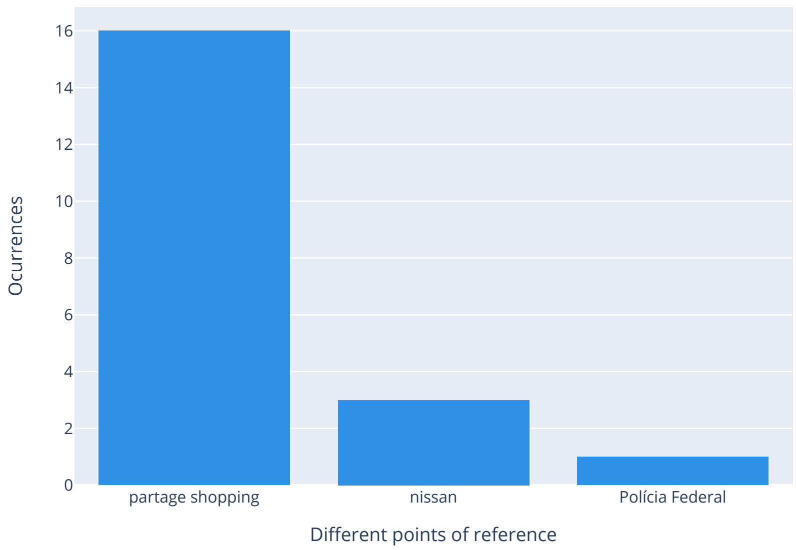


Figure 10: Further away places are used as reference more often than closer ones, as long as they are prominent.

Chapter 5

Spatial Relations Used by People in Conversation

After learning more about the usage of spatial relations by people, one can realize that the relationships most often used by people in dialogue are not present in the major spatial query languages. A natural next step was to try to implement them. This chapter proposes a set of algorithms that project in a 2D map, some of the relationships that are most often used by humans in daily communication. They take as input the geometry of the object that serves as reference in the spatial relation and return the polygon that corresponds to the real-world region described by it.

5.1 Spatial Relations Algorithms

As already mentioned in Section 4.2.3, the spatial relations “In front of”, “Near” and “Next” were by a large difference the most often used to describe locations in Experiment 1. Other than these, the three relations “At street”, “Between” and “Right of/Left of” were also chosen to compose a diverse set of relationships, for which the procedures were designed. The algorithms are presented as functions named as the spatial relations.

These algorithms must deal with some level of uncertainty when there is insufficient information about the spatial features referred to in the descriptions. For example, for a building located at a street corner, defining its facade may be considerably challenging or even impractical using traditional mapping data, posing even more challenges for modeling

Brazilian Portuguese Relation Name	English Translation
NA FRENTE DE	In front of
NA RUA	At Street
PERTO DE	Near
ENTRE	Between
AO LADO DE	Next to
À DIREITA DE	Right of

Table 5.1: Spatial Relations Names

some relations, such as In-Front-Of, as the buildings facade may be extended around the corner.

The lack of geographic data in the appropriate format may also impact the efficacy of this kind of algorithm. For example, in traditional mapping datasets many spatial extents of landmarks are available, however, many others are represented as single geographic coordinates (points). The algorithms proposed here are able to better infer the spatial extents of regions for polygon inputs. In the absence of this format of data they are also capable of working with point inputs. However, we believe the availability of this kind of data as polygons tends to increase considerably in the next years, contributing directly to the accuracy of systems that will incorporate those algorithms.

All proposed functions take as input a geometry, representing the spatial extent of a landmark, and return a generated polygon, representing the spatial extent of a region that best describes the relation with respect to that landmark. These regions are called here **acceptance regions**. An important observation is that the algorithms work in the scope of streets. This way, when generating the acceptance region to the Right-Of relation for instance, one should expect that the algorithm will produce a region that encompasses the portions of street that are to the right of the landmark.

5.1.1 In Front of

The intuition behind finding out the front of a landmark is that the streets that are closest are good candidates. After selecting the candidates (Figure 11), a simple verification if there is any other object between the landmark and the street candidate filters out the ones that

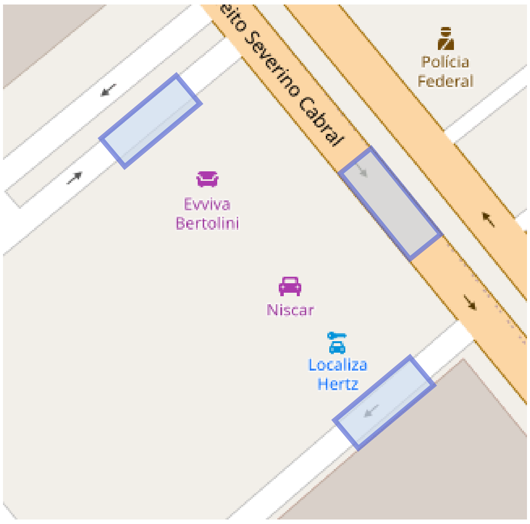


Figure 11: Street candidates for the front region of the landmark *Niscar*.

Figure 12: The buffer (represented by the blue square) that extends itself from *Niscar* to the street, intersects *Localiza Hertz*. Therefore, the street is in front of the latter, not the former.

probably should not be included in the final result since if this is the case, than that portion of the street is probably in front of the other object and not the landmark itself as depicted in Figure 12.

Algorithm 1 implements the relation In-Front-Of. Line 2 tests whether the input geometry is of type point or polygon. For point input geometries a buffer around the input is computed (Line 3), the intersection between this buffer and the nearby streets (Line 4) represents the candidates to be included in the acceptance region. For each candidate, the algorithm tests if there is another object between the input landmark and the candidate and includes the street in the final result, if it does not meet these conditions (Lines 5 to 10).

For polygon input landmarks, the procedure is almost the same. With the exception that the buffer used to select the candidate streets as well as the tests that check if a candidate street is really in front of the landmark (Figure 12), can be computed for each of the lines representing sides of the polygon (Lines 12 to 19). This allows the generation of an acceptance region that is much more accurate, for it really represents the full extension of the region that is in front of each particular side of landmark (Figure 14), as opposed to an estimate of such region, which is the case for point input landmarks (Figure 13). Line 12 returns the

Algorithm 1 In Front of

```
1: function INFRONTOF(landmark geometry)
2:     if landmark is of type point then
3:         Compute a buffer around landmark
4:         intStreets = the intersection between the buffer and all streets that intersect it
5:         for each street in intStreets do
6:             testLine = a line from landmark to street
7:             if testLine does not cross another landmark or street in intStreets then
8:                 finalFront = Union of street and finalFront
9:             end if
10:        end for
11:    else
12:        for each side in the landmark polygon do
13:            Compute a one-sided buffer in the line representing the side of the polygon
14:            streetFront = the union of all streets that intersect the one-sided buffer
15:            Compute a buffer between the landmark and streetFront
16:            if There are no other objects inside this buffer then
17:                finalFront = Union of streetFront and finalFront
18:            end if
19:        end for
20:    end if
21:    return finalFront
22: end function
```

acceptance region produced by the union of street candidates for the appropriate input format (Line 8 for points and Line 17 for polygons).

5.1.2 At Street

An example sentence that uses this relation is: *"The car is at the university's street"*. When the university is a well known landmark in the area, the street in which it is located becomes a common landmark. This relation produces an acceptance region that includes the whole

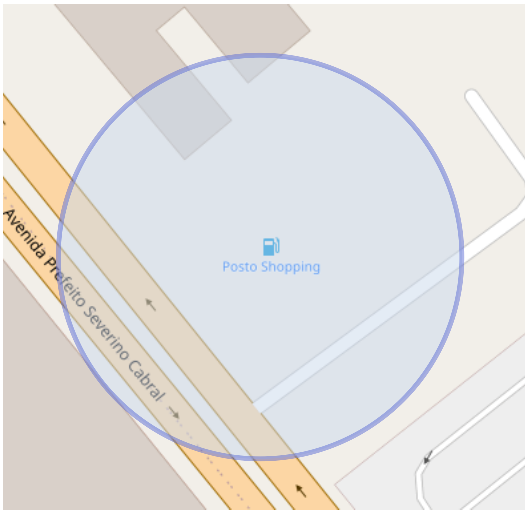


Figure 13: Buffer computed around the landmark when it is a point.

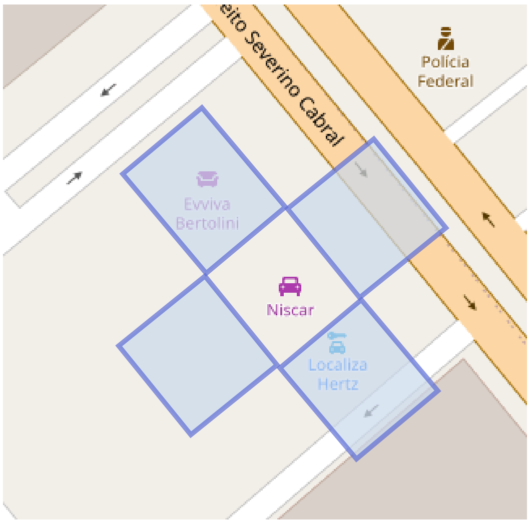


Figure 14: One-sided buffer, computed for each of the sides of the landmark polygon.

extension of the streets that intercept the front of the landmark.

Algorithm 2 At Street

1: **function** ATSTREET(landmark geometry)
2:     Use Algorithm 1 to compute the front of the landmark
3:     **for** Each street that intersects the landmark's front **do**
4:         **if** Area of intersection between the street and the front is bigger than some threshold **then**
5:             finalStreet = Union of intersection and finalStreet
6:         **end if**
7:     **end for**
8:     **return** finalStreet
9: **end function**

Algorithm 2 Line 2 computes the front region of the landmark by making use of Algorithm 1 as exemplified in Figure 15. Lines 3 to 7 includes in the acceptance region all streets that intersect the front as shown in Figure 16. For this relation it is important to filter out the parts of streets whose areas are small enough (line 4), as sometimes the crossing between streets is included in the front area (mostly for points) but only one of the streets is really in

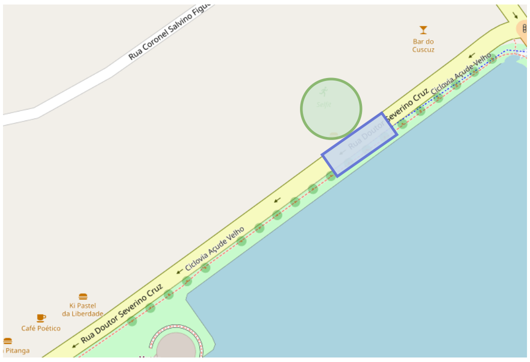


Figure 15: The blue rectangle represents the front of the green circle.

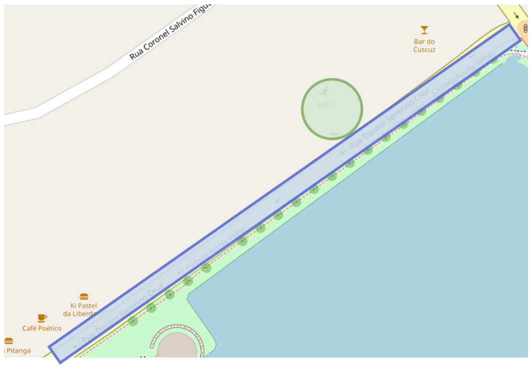


Figure 16: The entire street that intersect the front of the green circle.

front of the landmark.

If the data includes the addresses of the objects, the projection of the acceptance region could be thought as straightforward, however we also consider that people might think that streets that are not the official address of some building but that are adjacent to one of its sides could also be seen as "the building's street".

5.1.3 Near

The Near relation is implemented in Algorithm 3. It is quite simple and is the same for points and polygons. A buffer around the landmark, computed in Line 2, represents the region that is *near* it as in Figure 17.

Algorithm 3 Near

```
1: function NEAR(landmark geometry, distance float)
2:     return a buffer with a distance-sized radius around landmark
3: end function
```

The distance parameter should be tuned, and probably varies depending on the context (e.g. people who live in smaller cities might consider as near, a distance that is different from people that live in bigger cities). The analysis done in Section 4.2.4 can be used as basis to select a value to this parameter.

Figure 17: The blue polygon represents the region that is near the landmark pictured by the red polygon.

5.1.4 Between

Between is the only ternary relation in this list. It defines the position of one object, with respect to two others as in *"The car is between the university and the bookstore"*. For this reason, Algorithm 4 takes as input two geometry parameters.

Algorithm 4 Between

1: **function** BETWEEN(landmark1 geometry, landmark2 geometry)
2: Draw a line between a point in the surface of each of the two geometries
3: Get two points in the line that are at a distance d from each end
4: Draw a new line between the two points
5: **return** a d-radius buffer around the new line
6: **end function**

The main idea behind this algorithm is to draw a line between the two reference landmarks and compute a buffer around it. The function then starts by drawing this line (Line 2). However, if a buffer around this simple line is returned, as Figure 21 demonstrates, the result will include regions that are actually outside the desired relation. To fix this issue, the proce-

Figure 18: Both red circles are positioned from the same distance *d* to the nearest of the line ends.

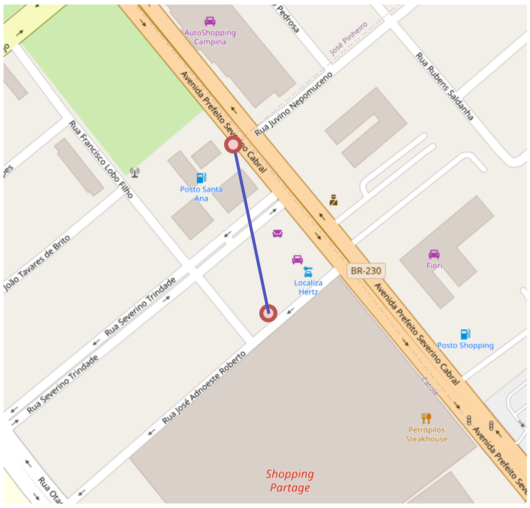


Figure 19: A new line is computed, now between the two red circles.

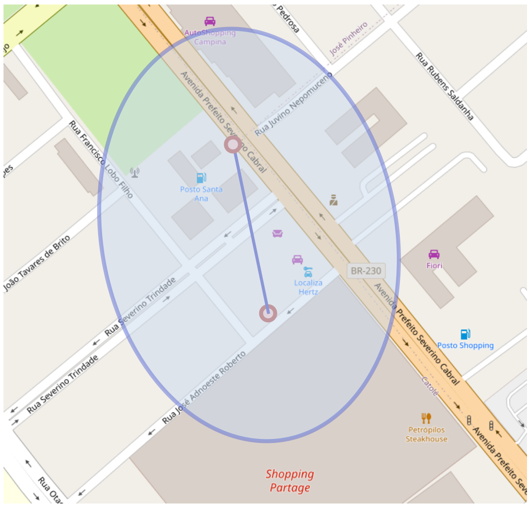


Figure 20: A buffer of radius *d* around the new line defines the acceptance region.

dure finds two points along the line that are located at the same distance *d* to each of the lines ends (Line 3, Figure 18). In the PostGIS [1] spatial extension for the PostgreSQL database, this could be done using the function ST_LineInterpolatePoint. A new line between these two points is drawn (Line 4, Figure 19). This new line is smaller than the original one and each of its ends is positioned at a distance *d* from the closest reference landmark. For this reason, a buffer of radius *d* around it (returned in Line 5) will produce a region that extends itself exactly from one landmark to the other, including no region behind them as shown in Figure 20.

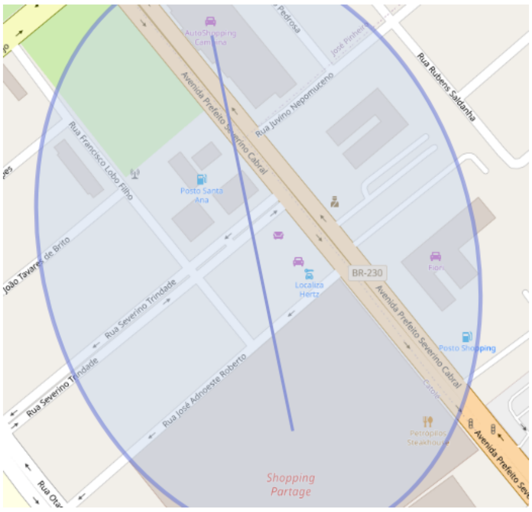


Figure 21: A buffer around the line between the landmarks will include regions behind them.

5.1.5 Next To

In the Next-To relation, regions that are immediately next to the landmark but not necessarily in front of it are included. The reasoning behind this algorithm is to compute the region that includes the streets that are near the landmark (with a smaller radius) and subtract from this region, the pieces of street that correspond to the front of the landmark. For this reason, Algorithm 5 starts by computing the landmark front (Line 2) using Algorithm 1. The street relation is also used, so Line 3 uses Algorithm 2 to compute the street relation. A buffer around the landmark is generated (line 4, Figure 22) and intersected with the region that correspond to the street of the landmark (Line 5, Figure 23). This intersection is then divided

[1] https://postgis.net/

in pieces that are separated by the crossing of streets. For each piece in this intersection (Lines 6 to 18) a line is drawn starting from the landmark (Line 7, Figure 24), if this line crosses the difference between intersections and the piece (Figure 25), this means that the intersection includes a piece of street that is closer to the landmark, so the one that is farther away is removed (lines 8 to 10). A special case is when the input geometry is actually in point format, for the front relation for points can include large regions. In this scenario, Lines 13 to 17 find for each street in the resulting area, the point that is closest to the input landmark. Buffers around these points serve as the front relation for the input landmark. After this, Line 19 returns the difference between the resulting area and the front as depicted in Figures 26 and 27.

Algorithm 5 Next To

```
1: function NextTo(landmark geometry)
2:     Compute landmark front
3:     Compute landmark street
4:     Compute a buffer around landmark
5:     nextInt = the intersection between the buffer and landmark street
6:     for Each partOfStreet that intersects nextInt do
7:         Draw a line from landmark to partOfStreet
8:         if Line does not cross the difference between nextInt and partOfStreet then
9:             nextFinal = union between nextFinal and partOfStreet
10:        end if
11:    end for
12:    if Landmark is of type point then
13:        for Each partOfStreet that intersects nextFinal do
14:            Get the point in partOfStreet that is closest to landmark
15:            nextFinal = nextFinal minus buffer around the closest point
16:        end for
17:        return nextFinal
18:    end if
19:    return Difference between nextFinal and the landmark front
20: end function
```

Figure 22: A buffer around the landmark and the output of the street relation.

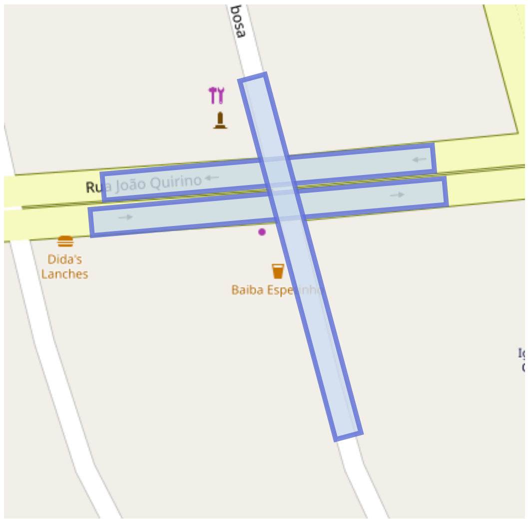


Figure 23: The intersection between the buffer and the street relation.

Figure 24: A line from landmark to a piece of street in the Near relation.

Figure 25: The line crosses the difference between the intersection and the piece of street.

5.1.6 Right of (Left of)

The Right-Of Algorithm 6 receives a string of text as second argument, representing the type of frame of reference that should be used to define the relation. It can assume the values of two of the three types mentioned in Section 2.3.1, intrinsic (defines right, based on the landmark front) and deictic (defines right based on the point of view of an observer

Figure 26: The pieces of street that are closer minus the output of the front relation (in red).

Figure 27: Result of the Next relation for the landmark *Baiba Espetinho*.

positioned in front of landmark and looking towards it, as in Figure 4).

The “Next to” relation produces regions that are positioned to the right and to the left of landmark. In order to find the region to the right, the algorithm tries to find the correct polygon from those that are considered to be next to, according to the appropriate frame of reference.

This function computes the front relation using Algorithm 1. For each of the streets that intersect the front acceptance region, (Lines 7 to 15) it computes a one-sided buffer on a line that goes from the input landmark to the street as depicted in Figure 28 (Lines 8 and 9). To determine in which side of the line the buffer is generated, the string representing the frame of reference is used (Lines 2 to 4). The relation Next-To is computed (Line 6) and if any of its containing polygons intersect the one-sided buffer, it is included in the final result (Figure 29). Line 16 returns the acceptance region, formed by the polygons of the Next-To relation that intersect the one-sided buffer.

The algorithm for the Left-Of relation is almost the same as this one, the only difference is in lines 2 and 3, where the “left” and “right” values are swapped. For brevity reasons, it is not included here.

Algorithm 6 Right of

```
function RIGHTOF(landmark1 geometry, for text)
    if for == "intrinsic" then bufferSide = "left"
    else if for == "deictic" then bufferSide = "right"
    end if
    Compute landmark front
    Compute landmark Next To relation
    for Each partOfStreet that intersects landmark front do
        Draw a line from landmark to the centroid of partOfStreet
        Create a one-sided buffer that grows in the direction of the bufferSide variable
        for Each polygon in landmark Next To relation do
            if polygon intersects buffer then
                finalRight = union between finalRight and polygon
            end if
        end for
    end for
    return finalRight
end function
```

5.2 Evaluating the Precision of the Algorithms

To grasp into the representations of spatial relations in the minds of people and to evaluate the precision of the algorithms, a second experiment was carried out. Volunteers were asked to read phrases containing references to landmarks and spatial relationships and then to draw on a map the polygon(s) they thought that best describe(s) the region referred to in the text. These geometries were used to assess how well the algorithms output represent the mental representations of people.

5.2.1 Research Questions

This experiment was designed in an attempt to answer two important research questions:

- When using spatial relations in conversation, what is the mental picture of the regions

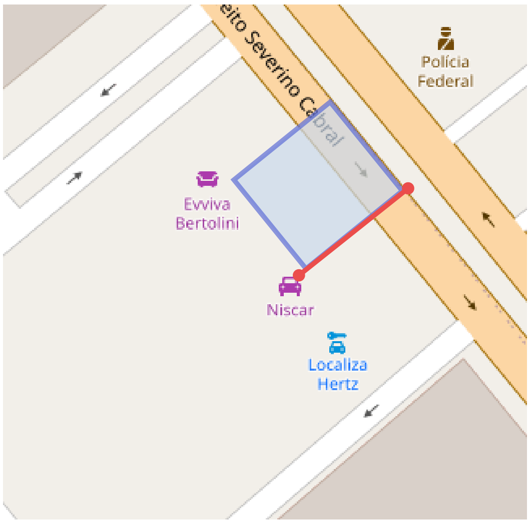


Figure 28: A one-sided buffer (in Blue) for a line that goes from the landmark to its front is computed considering a deictic frame of reference.

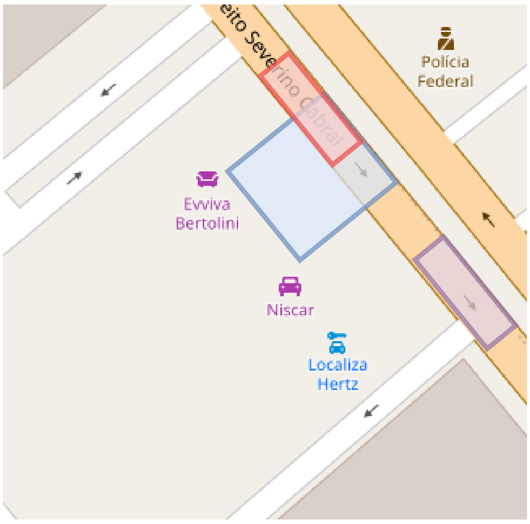


Figure 29: Only the red portion of the "next to" relation intersect the one-sided buffer in blue, therefore it is the only polygon included in the final result.

described by these relationships, that humans have in their minds?

- Are the output of the algorithms, good representations of these mental pictures?

5.2.2 Materials and Methods

Through the usage of a web app, another group of 20 participants, none of whom participated in Experiment 1, were told to picture the following scenario:

"Imagine that a friend will give you a ride and tell you over the phone where the car stopped and is waiting. Based on the description he gave you, we ask you to draw on the map the area where you think the car might be."

The participants of this experiment form a diverse group of people from different backgrounds. However, most of them are students (undergraduate and graduate) aged between 20 and 35.

The web app then shows up a map with a highlighted landmark and a sentence that describes the location of the car. Figure 30 shows the screen that the participants see when they are supposed to start drawing, with all the text translated to english. The sentence in (2) means *Your ride awaits you at: AT Café Poético's STREET, NEXT TO Bar do Cuscuz.*

The blue capitalized words represent spatial relations while the black ones represent spatial landmarks. Participants drew the regions by clicking on the map and creating points and lines. It is also possible to draw multiple disconnected geometries, to support scenarios where a participant wishes to draw on more than one place. Each person had to draw five relations for each of the four landmarks.

Figure 30: Translated interface of the web application used in the experiment

A street might extend itself for kilometers, and this could jeopardize the experiment, since participants could get tired of drawing really large areas. For this reason, relations At-Street and Next-To were combined so that participants were supposed to draw a polygon on only a smaller portion of the street.

The drawings were then stored in the GeoJSON format and a CSV of the data is available at GitHub [2].

5.2.3 Precision Evaluation

In order to evaluate the precision of the algorithms, they were implemented using PostGIS (The code is also available at GitHub[3]) and executed for each of the four landmarks used in the experiments. The polygons produced by them were then compared against the collected

[2]https://github.com/jslucassf/geoinfo-spatial-relations

[3]https://github.com/jslucassf/everyday-spatial-relations

drawings. One issue with the drawings was that although the experiment defined that participants should imagine the location of a car, some drawings do not intersect streets at all. This might be due to not so clear instructions and a future experiment can try to address this issue. However, as the algorithms function in the scope of streets (the regions produced by them are mostly located on the streets), the drawings that do not intersect streets at all were not considered.

Intersection of Areas

The chart presented in Figure 31 shows that for almost all relations, the algorithms produce regions that intersect the majority of drawings made by participants of the experiment. The relation Right-Of got the lowest results however this could be explained by the uncertainty that emerges from the ambiguous nature of this relation.

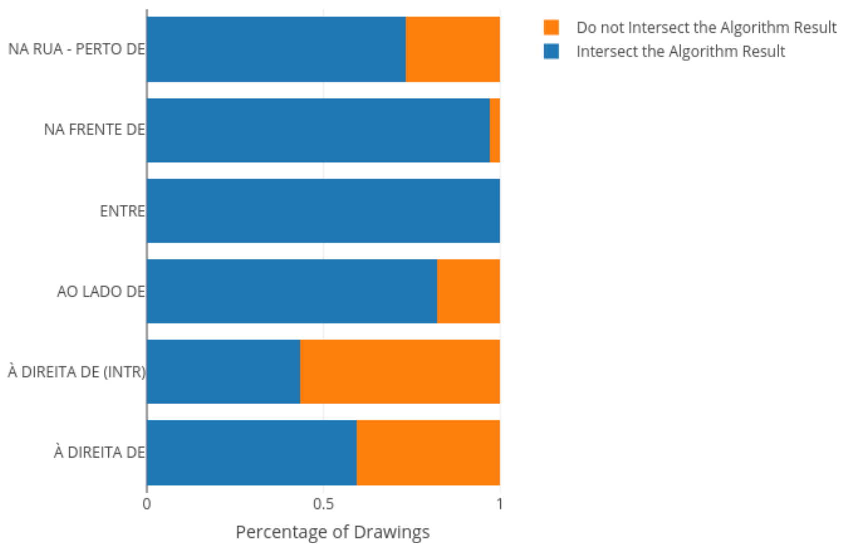


Figure 31: Percentage of drawings that intersect the region produced by the proposed algorithms.

Jaccard's Similarity Coefficient

A common metric used to access the similarity between sets is Jaccard's Similarity Coefficient. It expresses how similar two sets are in a scale of 0 to 1 and is computed by the Equation 5.1. This metric was used to evaluate how similar are the geometries produced by the algorithms and the drawings made by the participants.

$$Jaccard(A, B) = A \cap B / A \cup B \tag{5.1}$$

In order to assess the complexity of the task, a value to show how similar the drawings made by the participant's are with each other was also computed, here it was called the **inner jaccard**. For each drawing, the Jaccard's Similarity Index with all other drawings in the same category (landmark and spatial relation) is computed, the median result is the inner jaccard and it represents how similar is this drawing to all the others. Figure 32 displays the results of the analysis.

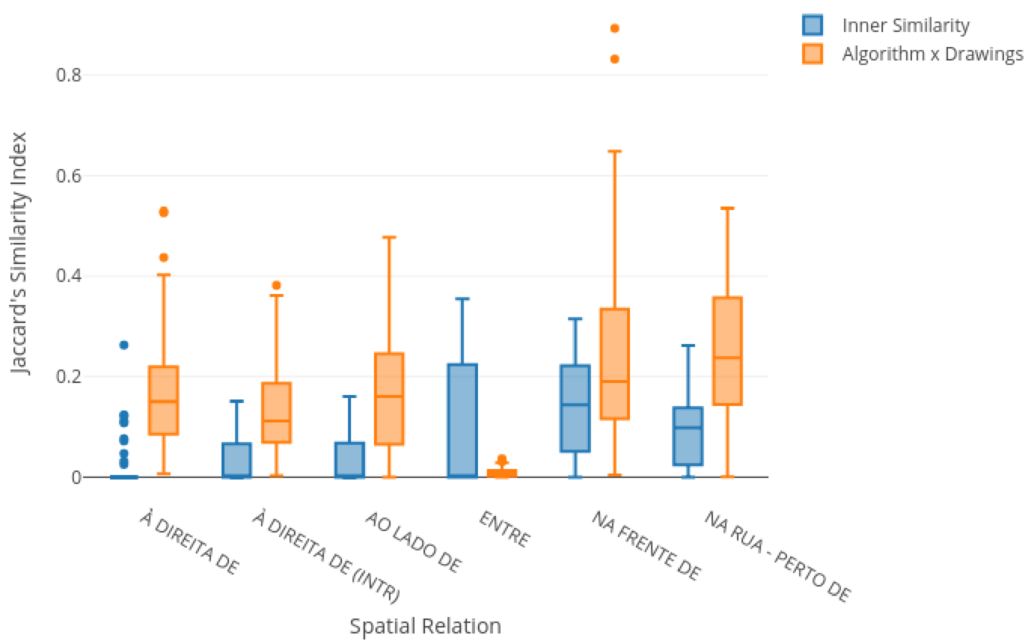


Figure 32: Jaccard's Similarity Coefficient between geometries

As can be seen in the low inner jaccard values, the drawings themselves are not very similar. This might indicate that people have different understandings of spatial relations. Considering this, the proposed algorithms had modest results when compared to such a diverse set of polygons. These results, when coupled with the high intersection percentages shown in Section 5.2.3, suggest that these algorithms are a good starting point for the implementation of some of the spatial relations that are most used in the daily language of people.

The next chapter proposes a conceptual model of conversations about the location of objects in an urban space. This model makes use of some of the concepts explored so far, such as spatial relations and spatial landmarks.

Chapter 6

Modeling Spatial Dialogue

Trying to impose structure upon an unpredictable event such as a human conversation might prove itself to be a rather challenging task. However, having an accurate model of a closed scope version of such an event might be an achievable feat. A conceptual model can capture the most important entities, attributes and relationships involved in these conversations, and thus assist in the development of computational dialogue systems that work in the specific domain, being used in the specification stage in the project of such systems.

In this chapter one instance of the aforementioned model is proposed. It specifically tries to represent a dialogue in which a person tries to convey the location of an object in space whose whereabouts is of mutual interest to all of the involved stakeholders. As this type of dialogue involves complicated steps, another conceptual model is also proposed, representing one specific concept that requires its own concepts and relationships.

6.1 Concepts

A conceptual model work as a mental conceptualization obtained through the observation of the real-world event. Concepts are the most important parts of these representations and describing them facilitates the understanding of the system being modeled. This section lists and explores all the concepts that participate in the process of a location description conversation.

6.1.1 Person

Represents any person that takes part in this specific type of conversation. A minimum of two are necessary for a conversation to take place. Considering how many people could be involved in such a discussion, one could realize that the term "conversation" is in fact vague. It is known that two people can constitute a conversation, and certainly also three as well. At which number this meeting stops being a conversation? Here it is assumed that a conversation can be established by at minimum two and a maximum of n people.

Possible instances of the concept person can be represented by names such as *Mike*, *Jimmy*, *Kim* and *Chuck*.

6.1.2 Message

People communicate through an exchange of messages. These are thus, the fundamental units of conversations. Messages are composed of words, grouped together in a way to convey meaning and useful information. Delivered through many mediums such as through the air in the form of speech or through written word, in this specific domain of dialogue, messages can come in many types.

Location Description

Messages whose sole purpose is to describe where a specific object is located in space. This type of message usually makes use of other important concepts: spatial relations and references to landmarks. Location descriptions have been studied and classified in two different types: Destination and Route descriptions [44]. Destination descriptions locate places based on reference points in the vicinity. Route descriptions are composed of step-by-step instructions to reach the place starting from an initial point. While some descriptions can still combine characteristics from both categories, it is postulated that destination descriptions are usually shorter, therefore the cognitive workload of producing them is smaller.

Possible instances of location descriptions:

- *"I am in front of the bookstore, near the city hall."*.
- *"Turn right after the semaphore and go ahead until you see a big old church. The place you are looking for will be located to your left."*

Disambiguation Question

Issued when a previous message includes ambiguous expressions, disambiguation questions are really important to the flow of conversation, since they help us cope with the uncertain essence of natural languages. The ambiguous bit of the message can be a reference to a landmark or about the nature of the mentioned spatial relation.

Possible instances of disambiguation questions:

- *"There is more than one bookstore in the vicinity, to which one are you referring to?"*.
- *"When you say to the right of the store, you mean to the right side from my point of view or the the right side of the store facade?"*

Denial / Confirmation of Knowledge

Oftentimes, before producing a location description, the message issuer can inquiry the recipient about whether he knows a particular landmark or not. This type of message helps speed up conversation for it allows the person describing a location to tailor a description to the knowledge of the audience. This facilitates the interpretation of the message. The recipient on its turn responds either by denying or confirming knowing the place.

Possible instances of Denial / Confirmation of Knowledge:

- *"- Sender: Do you know that bookstore near the city hall? - Recipient: Yeah I do!"*.
- *"- Sender: Have you ever been to that clothing store near the big old church? - Recipient: No I have not."*
- *"- Sender: It is located right in front of the city hall. - Recipient: I'm sorry but I'm new to the city and I don't know where the city hall is located."*

Further Inquiry

After receiving a location description some doubts about the location can continue to exist. The described region might still be too large or the exact location is yet not clear. In these cases, the recipient of the description can ask for more information.

Possible instances of Further Inquiry:

- *"Ok, lots of places are near the city hall. Can you be more specific?".*
- *"I understand that you're talking about the church's street. But in which side exactly?"*

A further inquiry might even include spatial relations and references to landmarks. In this case, the roles are reversed and the recipient of the location description is the one producing a location description. As an example: "Ok, lots of places are near the city hall. Is it closer to the bookstore?"

6.1.3 Description Interpretation

Upon receiving a location description, the message recipient has to interpret it. This is one of the most important activities during this event. As was already mentioned, such a message contains references to landmarks and spatial relations. The recipient identifies the references and projects in his mind, the region being described by the sender of the message. The output of this activity is the mental representation of the region that results from the application of the spatial relation and the referenced landmarks.

6.1.4 Acceptance Region

The mental picture of the region that might contain the object whose location is the main target of the conversation. In an ideal scenario it should be the same every time the same spatial relation is applied to the same landmarks. However, it is the product of the interpretation of a message by a person, therefore its shape is subject to the semantic of the expressions chosen by the conversation participants, as well as their knowledge of the city.

Whenever an acceptance region is produced, the person who interpreted the message evaluates if it has enough information about the location. As already mentioned, the region might still be too large and thus, additional inquiries may be necessary.

6.2 Dialogue Conceptual Model

A diagram of the conceived conceptual model for this dialogue domain is presented in Figure 33. It tries to capture the main concepts involved in the dialogue process that have been explored in the previous section. An important thing to notice is that this is the model of a dialogue that can include two or more participants as can be seen in the cardinality of the message to person relationship. This is due to the fact that it is entirely possible for someone to describe the location of an object to more than one person. The Dialogue process happens through the involved parties sending and receiving various types of messages and interpreting them. As this is a location description scenario, interpreting one of such descriptions produces a mental representation of a region that best matches the description.

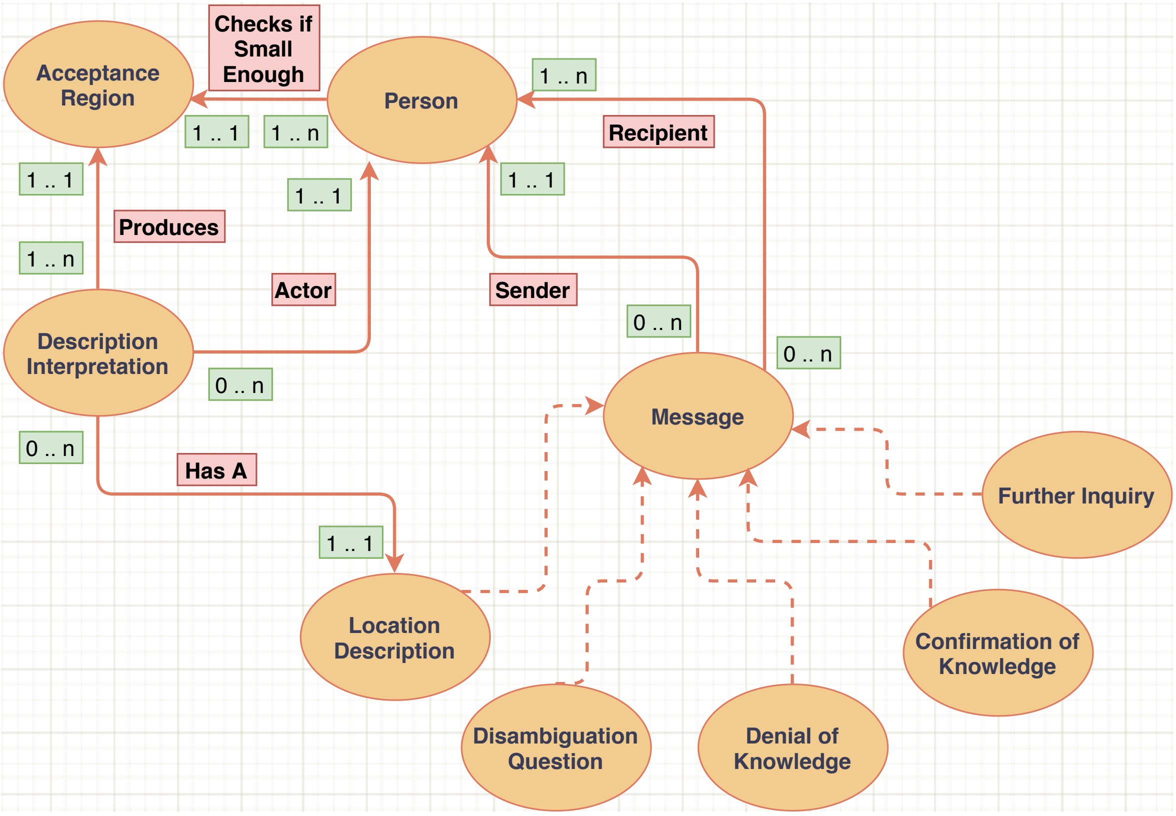


Figure 33: Dialogue Conceptual Model

As already mentioned in Section 6.1.3, the *description interpretation* concept is probably one of the most important and complex stages of the dialogue process. So much that it deserves to be zoomed in and analyzed. The next section proposes a second conceptual model, specific to this sub task.

6.3 More Concepts

This second conceptual model has its own constituent concepts. As some of them have already been defined, in order to avoid repetition, the concepts *message* and *acceptance region* will not be explored. To understand them and read a few examples, refer back to Sections 6.1.2 and 6.1.4.

6.3.1 Reference to Spatial Relation

There are several ways to refer to a spatial relation. When someone wants to state that an object is located near the city hall, this spatial relationship could be described in many different ways. For this reason, a crucial step in the interpretation process, would be to map to which specific spatial relation, a particular expression is referring to.

Possible instances of References to the Spatial Relation "near":

- *"near"*
- *"close to"*
- *"nearby"*
- *"in the vicinity"*

6.3.2 Spatial Relation

A spatial relation defines the position of an object in space, in relation to a reference object. A good list that serves as a starting point for possible relations can be seen at Table 4.1.

The distinction needs to be made from the references to relations, explored in the previous section, to the abstract idea of the spatial relations themselves. Each reference can be mapped to a single relation.

6.3.3 Landmark Alias

In a similar way to spatial relations and references to spatial relations, when people speak, they tend to use aliases to refer to specific places. Funny nicknames, historical terms and even shortenings of the original names of the places. Mapping each name to a specific place can be a difficult task, but it is essential nonetheless, since aliases can be responsible for a lot of confusion in communication as they can be ambiguous.

Possible instances of Landmark Aliases:

- *"Big old church"*
- *"The city of love (In reference to Paris)"*

6.3.4 Landmark

A landmark is a recognizable place, natural or artificial feature that is often used as reference (through the usage of an alias) to describe a location. A landmark can be a building, a body of water or even man-made monuments.

Possible instances of Landmarks:

- *"The actual church"*
- *"A gas station"*
- *"The Eiffel Tower"*

6.4 Description Interpretation Model

The conceptual model for the event of interpreting a location description is presented in Figure 34. A message is usually composed of references to landmarks and spatial relations. There are different linguistic expressions to describe these components. For this reason, an important step in the interpretation process is to match the expression to the appropriate real world landmark and the abstract idea of one particular spatial relation. The spatial relation produces an acceptance region, by modifying the spatial extent of the associated landmark. Figure 34 follows the same notation conventions described in Section 2.1.1.

6.5 Factoring in Time

A conceptual model is a powerful tool that can capture the fundamental concepts of a phenomena that takes place in the physical world. However, in a dialogue domain the aspect of

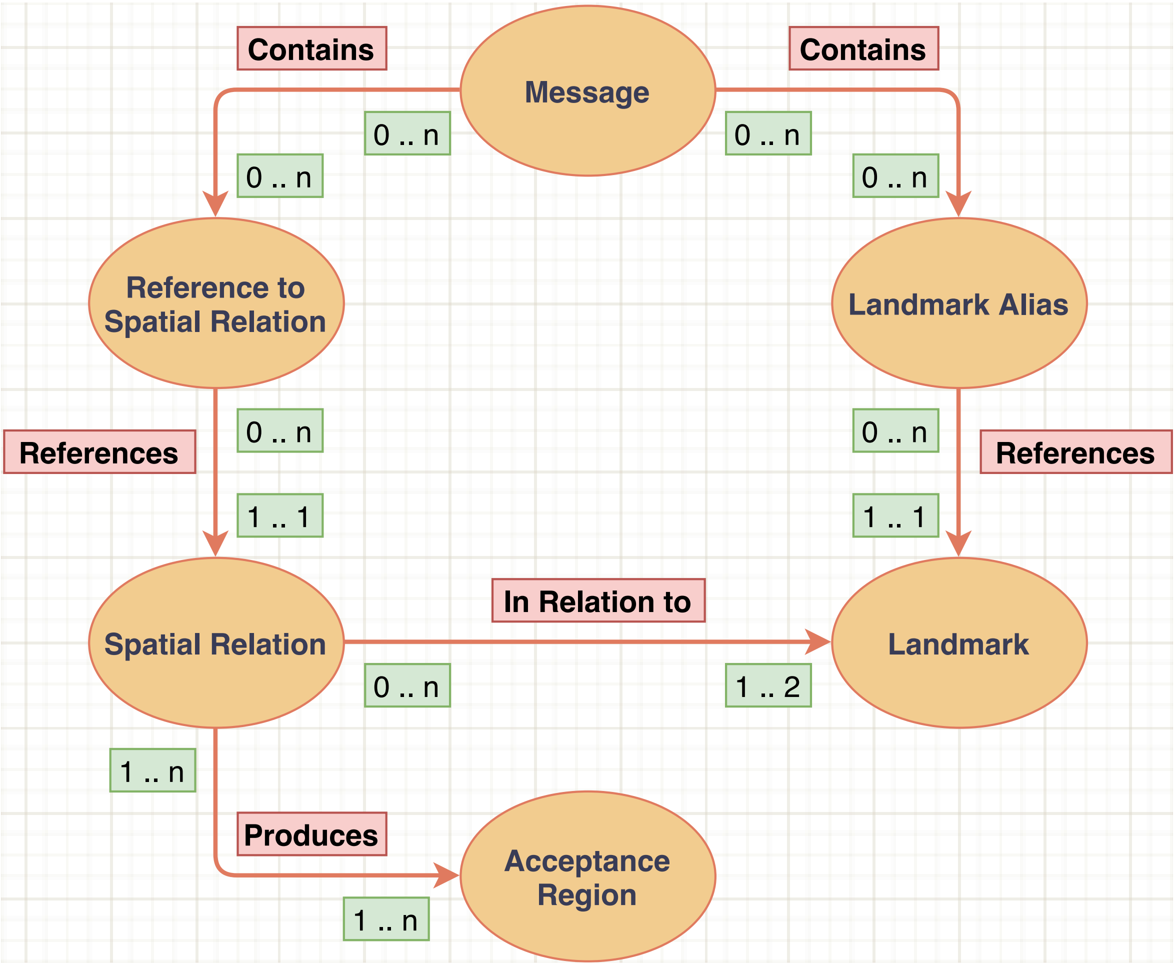


Figure 34: Location Description Interpretation Model

time is crucial. Knowing what message comes before and after can enhance the understanding of the system being described.

The proposed models can be used to assist in the development of computational dialogue systems, also known as chatbots. To include the time factor and illustrate the usage of the models in a scenario where a bot assumes the role of the person to whom the location is being explained, Figure 35 presents an activity diagram that demonstrates the flow of a conversation that follows the proposed models.

The conversation starts with the bot receiving a location description, and then it tries to identify the landmark references included. If some level of ambiguity is detected, it asks a question back to the person. The process continues with the bot checking whether it has a representation of the given landmarks in its database. If it does not, more information is needed. Once the bot has access to a description with unambiguous references to known

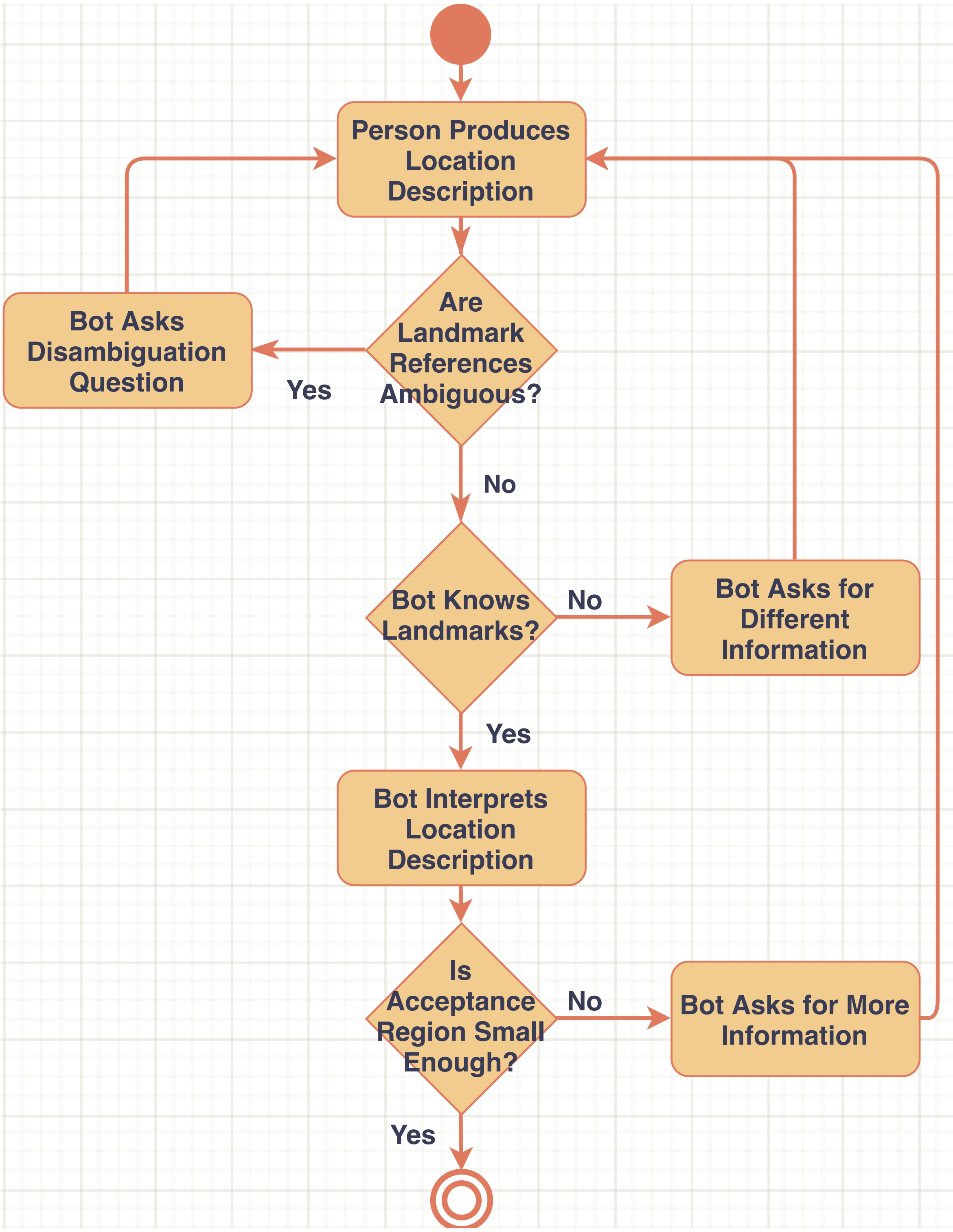


Figure 35: Activity Diagram of a Dialogue

landmarks, it projects the spatial relations and produces an acceptance region. If the region is small enough, the conversation ends. Otherwise, the bot asks for additional information, in order to restrict even more the resulting region.

In the next chapter, a first step in the direction of the implementation of geographic aware conversational agents is taken. It describes the process and challenges of implementing an agent that follows the conceptual model described in Section 6.4 as specification and is capable of interpreting location descriptions.

Chapter 7

Implementing the Location Description Interpretation Model

The conceptual models presented in Chapter 6 are intended to assist in the development of conversational agents that are capable of acting in the problem of locating a place in an urban scenario. In this chapter, a dialogue system that works based on the concepts in the model that represents the location description interpretation process is presented. By extracting the required features from the text and executing the activities listed in Section 6.4, this conversational system can process a location description and produce a polygon that represents the location being described.

7.1 Scope

During the development of the agent, a few decisions had to be made. The landmarks detected in the text have to be matched to a real world spatial object. By virtue of the knowledge of places of the researchers and availability of test participants, a database containing geometries that represent these objects was created using data from the city of Campina Grande in the state of Paraíba, Brazil. For similar reasons, the agent was trained with textual messages in the Brazilian Portuguese language, therefore it produces and interprets messages in this language.

The conceptual models proposed in this thesis represent spoken location description dialogue. For this reason, even though textual data was used during the training stage of the involved models, these messages try to mimic the way people speak in daily conversation. As a result of this decision, the data does not contain common mannerisms that are often present in written exchanges such as shortening of words, be it in common expressions (e.g. using “u” instead of “you”) or in abbreviations to place names.

Concerning the supported spatial relationships, as algorithms were already implemented for the experiments described in Section 5.2.3, the list of supported spatial relations is the same of the proposed algorithms:

- In Front of
- At Street
- Near
- Between
- Next to
- Right of
- Left of

7.2 Spatial Relations and Landmark References

As already mentioned, when people produce location descriptions references to spatial relations and named places are usually produced and theses references can assume many forms. The first step in the description interpretation process is to identify these references in the text. As an example of the task, a location description such as *“In the same street as the coffee shop”* contains the spatial relation reference “In the same street” and the landmark reference “the coffee shop”.

In order to perform this task, once again a machine learning classifier based on Conditional Random Fields (CRF) [26] was used. The model was trained to identify 8 classes of ngrams, these classes are the seven spatial relations listed in the previous section (named “sr_front”, “sr_at_street”, “sr_near”, “sr_between”, “sr_next”, “sr_right” and “sr_left”) and finally, the “landmark” class. Therefore, having as input the previously mentioned location description, the model should label the expression "In the same street" as belonging to the class sr_at_street and "coffe shop" to the landmark class.

To train the model, the RASA NLU (Natural Language Understanding) open source library[1] was used. This library is commonly used in Named Entity Recognition tasks and for intent classification of messages. RASA uses a tensorflow[2] pipeline for training the CRF models. CRFs perform well as opposed to pretrained classifiers such as the ones in spaCy[3] since the goal is extracting custom entities that are relevant to the context at hand. They are able to generalize training data and, according to RASA NLU documentation[4], are effective with training data that includes at least twenty examples per entity. Thus, a training dataset with 150 sentences describing places was manually created and annotated. The data contains 191 examples of landmark references and 170 occurrences of spatial relations. To evaluate the model performance, testing was done using a subset of 57 sentences produced by the participants of the experiment described in Chapter 4 which correspond to approximately 10% of the sentences in that experiment. This subset was also manually annotated by the researchers. The model achieved a sensitivity score of 0.9.

7.3 Landmark Aliases and Matching

Matching a textual landmark reference to a real world spatial object is not an easy task. First, a database of the possible landmarks in the target location is needed. This dataset was built using data extracted from OSM (Open Street Maps). OSM Data includes geometries in vector format, representing buildings, roads, and even natural features such as bodies of water and trees.

Each record in this database contains a name field that can be used by a human to identify the spatial feature. However, this field usually contains official names and these are not always the ones that people use in spoken conversation. To illustrate this, let us consider the Federal University of the city. In the database, the record that correspond to the university has a name field value of "UFCG - Universidade Federal de Campina Grande" it is common

[1] `https://rasa.com/docs/rasa/nlu/about/`
[2] https://www.tensorflow.org/
[3] https://spacy.io/
[4] https://blog.rasa.com/rasa-nlu-in-depth-part-2-entity-recognition/

knowledge that while in conversation, people usually refer to it as simply "UFCG" or even "Federal". Another case of different name usage is when people refer to places using old historical names or even funny colloquial nicknames.

In light of all these complexities, a pre-processing is needed. A similarity measure was computed based on the Levenshtein[5] string distance metric. A list of the names of places was produced and the value of the expression extracted from the location description is compared against this collection. The record that ranks the highest in the metric is then selected. This similarity value accounts for some level of typing errors, and small variations in the names. For instance, a landmark reference of value "Partage Shopping" is matched to a record in the database named "Shopping Partage" with a similarity value of 95 in a scale of 0 to 100. Because of time and language constraints, the Levenshtein distance was chosen as a distance metric. As the intention is to model spoken conversation, a future work could try to make use of Phonetic Algorithms such as the Soundex[6]. This type of metric has the potential to generate more accurate results and with better time performance.

To address the usage of vernacular place names (e.g. historical and colloquial), the list of aliases is enhanced with the most common names used by people. This is a hard and time consuming task that could be benefited from automatic alias generation strategies. Having a small scope and admitting some level of failure, this step was done manually, producing a table that includes more than one record for each real world landmark that possess more than one commonly used name.

7.4 Dialogue Management

The main goal of the proposed conceptual models is to represent a dialogue, therefore the aforementioned classifiers are integrated with two other machine learning models, the first classifies the **intent** of the message while the second predicts the next **action** in the conversation.

Each message has an intent which represents the intention of the person who produced it. Intents can assume different values such as the sub concepts mentioned in Section 6.1.2.

[5] https://en.wikipedia.org/wiki/Levenshtein_distance
[6] https://en.wikipedia.org/wiki/Soundex

For this agent, messages can be classified in the following intents:

- Greeting
- Location Description

This means that the agent is ready to interpret salutation messages and location descriptions. On its turn of the conversation, the agent executes an action. Actions can include complicated functions such as projecting the geometry resulting from spatial relationships and landmarks or simply issuing plain utterances back to the user. The possible actions for the agent are:

- Describe Instructions
- Interpret Spatial Relation
- Denial of Knowledge
- Report Success

The conversational agent is thus capable of describing what is its purpose, interpreting a given spatial relation, informing to the user that it could not find the mentioned landmark and reporting the success of the location description interpretation process.

With the help of the RASA open source framework, the dialogue management model is trained using a type of data called stories. A story is a possible conversation example in the format of a sequence of intents and their appropriate actions performed in response. Stories are written in YAML[7] format and an example can be seen in Figure 36.

7.5 User Testing

A crucial step in the development of any conversational system is testing it with users after a decent version is available. This step allows the developer to peek into the way different users talk to the assistant. This is important for the developer can be focused on one particular type of message that not always represent the way people talk.

[7]https://en.wikipedia.org/wiki/YAML

```
version: "2.0"
stories:
- story: happy path
  steps:
  - intent: greet
  - action: utter_instructions
  - intent: place_description
  - action: action_report_success
```

Figure 36: Example Story of a Common Conversation

After having a simple version of the agent ready, 6 people were invited to give it a try. During these tests, the agent presented the user with a link to a location on Google Street View. It then asked the user to provide a location description to this particular spot.

After collecting and analyzing conversation data, one thing was made clear. People not always provide location descriptions that include only one spatial relation. Some people prefer to write a big description containing multiple spatial relationships and landmark references as depicted in Figures 37 and 38. This finding triggered many changes to the training data as it basically contained only descriptions with one spatial relation.

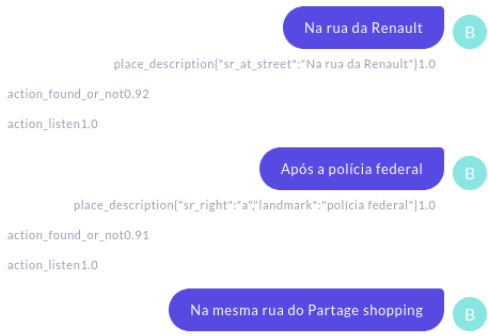


Figure 37: Multiple messages containing a single spatial relation each.

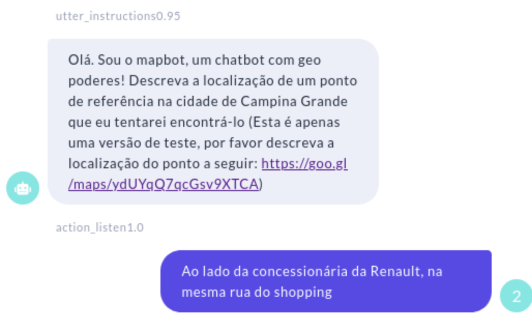


Figure 38: Multiple spatial relations in a single message.

7.6 Acceptance Regions

After the identification of the spatial relationships and landmark references in the description, after retrieving from the database the geometry data that represents the spatial extent of the landmark being referenced, an action that runs the appropriate spatial relation function is triggered. A polygon representing the interpretation of the location description is generated. Figure 39 shows the geometry generated after processing the message “Na frente da Korpus”, which means “In front of Korpus”.

Figure 39: Region in front of “Korpus Academia”

Figure 40 portrays the region that represents the spatial extent described by the sentence “Entre o Parque da Criança e o Partage Shopping”, Portuguese for “Between Parque da Criança and Partage Shopping”.

Following the discoveries from the user tests, the agent is capable of processing multiple spatial relations in a single message. First it identifies all spatial relations and the associated landmarks, then it produces acceptance regions for each relation and finally the intersection between these regions is returned.

Figures 41 and 42 demonstrate this process. The region in front of the landmark “AutoShopping Campina” is uncertain since it can be present in three different streets (Figure 41). By providing more information such as another spatial relation in the message, a person can help filter the relevant region, as was done in Figure 42 with the addition of an “At Street” relation. The full location description interpreted in Figure 42 is “Na frente do AutoShopping Campina, na rua do Shopping Partage”, that can be translated to “In front of AutoShopping Campina, at Shopping Partage’s Street”.

Figure 40: Region between "Partage Shopping" and "Parque da Criança"

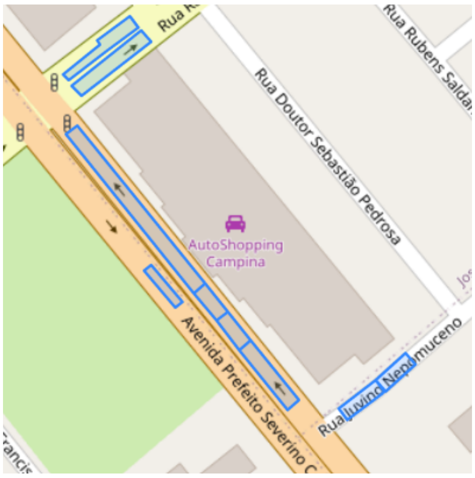


Figure 41: Region in front of the referenced landmark, "AutoShopping Campina"

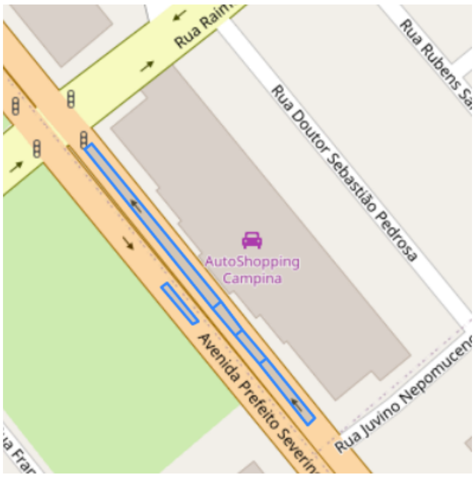


Figure 42: Region in front of "AutoShopping Campina" and at the same street as "Shopping Partage"

7.7 Architecture

The software architecture of the implemented conversational system is presented in Figure 43. A user enters a message to the bot server. This message is classified by the intent classifier and has its entities extracted by the entity extraction model and a landmark matching module. In possession of the appropriate message intent, the Dialogue manager model predicts the next action that the agent should take and calls an action server that executes it. Finally, the action server accesses a geographic database server to return a response back to the user.

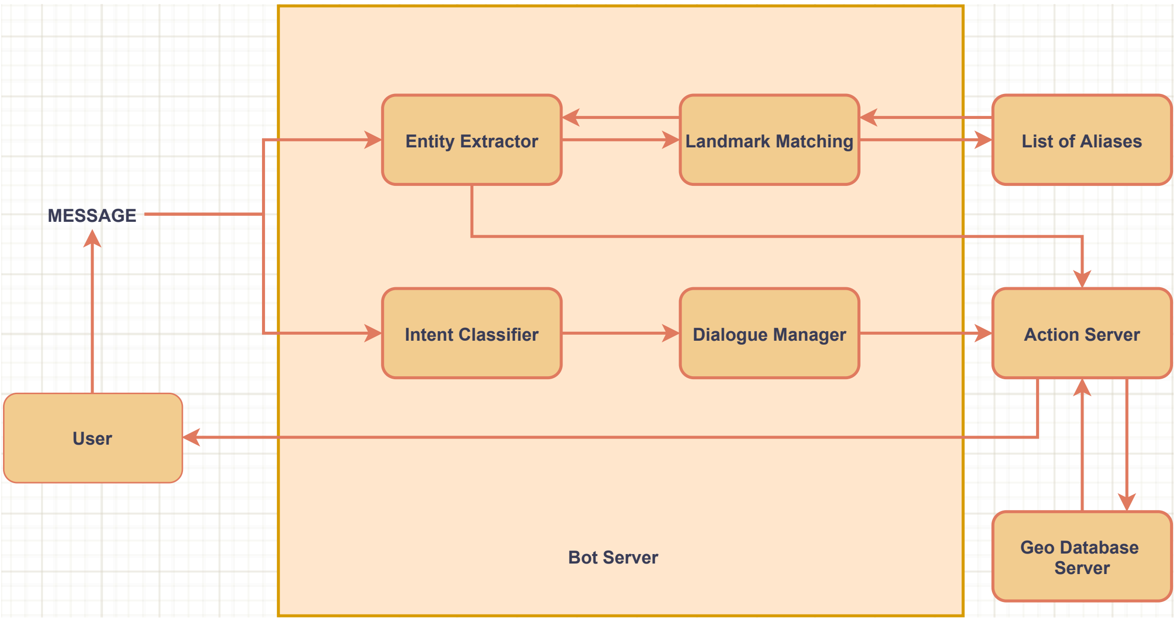


Figure 43: Architecture of the conversational system

Finally, the next chapter includes a brief discussion of the entire research, mentioning the problems that might affect it and also pointing lines of work that could be followed in the future.

Chapter 8

Conclusions

In this chapter, a brief discussion of the research is given. A section is dedicated to the limitations of the propositions in this work. Finally, a few pointers for subsequent work that can be done in the future are presented.

8.1 Discussion

The main objective of this thesis was to propose a conceptual model to represent real-world conversations where a person is trying to convey the location of an object in space to others. In the process of developing such a model several other studies and analysis have been made. For this reason, this work includes contributions to: the understanding of the way people reason about space (and, in particular, spatial relations); as well as to the modeling of qualitative spatial relations in conventional spatial query languages.

8.1.1 Spatial Reasoning and Spatial Relationships

Concerning spatial reasoning, one of the most prevalent semantic aspects of location descriptions seems to be references to landmarks, together with the usage of expressions that represent spatial relationships. These features are present in the majority of the descriptions that were analyzed and have the potential to present some insight into the way people reason about space.

Through inspection of the landmarks that participants chose to reference, it was possible to note that proximity to the location being described was not the decisive factor in the choice of the most appropriate landmark for descriptions. There seems to exist a correlation between the landmark choice and how well known or prominent these spatial features are. This process could be due to the idea that when trying to convey location information, people tend to try to tap into the knowledge of the environment that the addressee of the message possess. As the list of places a different person knows is seldom available knowledge, resorting to prominent places in the the environment might be a good strategy for reference selection because of the higher chances of that place being known by the other person.

The spatial relations that have been used more often were by a large difference "In-Front-Of", "Near" and "Next-To", despite the ambiguous nature of the relation "Next-To" and the vagueness of the relation "Near". Regarding this vagueness, through an analysis of the distances from landmark to goal location in descriptions using the relationship "Near", an effort was made to try to draw a line for the applicability of this term. In the studied data, most of the times, the distances referred to as being "Near" corresponded to a maximum of 300 meters. The medium "Near" distance, is then estimated to lie in the range of 144 to 183 meters.

The spatial relations used by people in the analyzed data are not always included in the list of relations supported by traditional spatial query languages. This motivated the design of algorithms to derive the geometries of some of the spatial relationships, when associated with reference landmarks. The algorithms cover the relations "In-Front-Of", "Near", "At-Street", "Between", "Next" and "Right-Of / Left-Of". An experiment that evaluates how well the output of the algorithms matches the mental representation of spatial relations in the minds of the participants was conducted. The analysis of the collected data shows that this is a difficult problem, however, the proposed algorithms hold promissory results, intersecting most of the regions drawn by participants and presenting some similarities to them. This experiment represents another contribution to the field, as it makes available a dataset of more than 400 drawings of spatial relations, allowing further studies on their interpretation by humans.

8.1.2 Conceptual Modeling

In order to assist in the development of conversational agents with spatial abilities, a conceptual model that represents dialogues in which a location is being described was proposed. Such theoretical representation depict relationships between the following important concepts that are present in communication: “Person”, “Message”, “Description” Interpretation” and “Acceptance Region”. Messages can be of different types, including “Location Description”, “Disambiguation Question”, ’Denial or Confirmation of Knowledge” and “Further Inquiry”.

The description interpretation is a complex process in its own, and thus it received special attention. Another conceptual model was developed, to represent this particular task. It includes the concepts “Reference to Spatial Relation”, ”Spatial Relation”, “Landmark Alias” and “Landmark”. The interpretation process produces a polygon that portray the region defined by the landmark used as reference and the spatial relation that modify its spatial extent.

These two models, combined, have the potential to improve understanding and aid developers in developing conversational agents that can tackle the task of locating places in an urban scenario, through conversation with humans. In fact, a chatbot that serves as proof of concept of one of the models has been developed. It is capable of classifying messages in the appropriate types and performing the interpretation of the ones of type “Location Description” according to the second model, producing an acceptance region that tries to match the image in the mind of the person that conceives a description. This interpretation makes use of the spatial relation algorithms presented in Chapter 5 and their implementation is influenced by the findings of the analysis in Chapter 4.

8.2 Threats to Validity and Future Work

The validity of the research described in this thesis can be affected by several factors. First and foremost, the experiments that were conducted had the attendance of Brazilian Portuguese speakers, more specifically, people from the northeast region of Brazil. As dialects can present differences between regions and even states, this variability can hinder the potential for generalization of an analysis of data produced by people from one specific background. A future work could try to reproduce the studies, using a dataset of descriptions produced by participants from a wider background of places, languages and cultures.

The findings related to the relevance of places suffer from the lack of a clear metric that divide landmarks in terms of “well known” or even “important” locations to a specific region. The definition of such measures of the prominence of landmarks, can allow a more robust investigation on the cognitive process of choosing references to compose location descriptions.

For a lack of available time, only a subset of the spatial relations detected in the description data were studied and implemented. In a real-world conversation scenario, people might describe a location making use of different spatial relations and even different types of descriptions (e.g. route descriptions). This limitation may frustrate the users of the chatbot as some of their messages certainly could be misinterpreted. In future studies, the design and implementation of algorithms for more spatial relations such as “Behind”, or even motion related relationships such as “Before” and “After” can improve the performance of the agent.

The availability of geographic data in the appropriate format is still an issue for many locations in the planet. Although this might change in the next few years, many landmarks are still represented in geographical databases as single geographic coordinates. A future study could try to make use of vector calculus to infer a generic polygon according to the angles of the surrounding spatial features, such as streets and other landmarks.

Another possible enhancement to a conversational agent system architecture would be to integrate natural language generation models for the interactive generation of messages. This advancement has the potential of transforming the dialogue into a more natural and fluid experience.

It has been detected that a few times when describing locations, people reference visual features of landmarks, as in “Near a house with a big red door”. In the present moment, conversational agents have no means to make use of this type of information to aid in the location process. In the future, with better spatial data that perhaps includes real pictures of places or even making use of satellite imagery, computer vision models could aid in the interpretation of such visual cues.

Finally, the chatbot implemented in this thesis only covers a portion of the proposed conceptual models, with more time available in a future work, a new version of this conversational agent could be developed. This agent should be able to issue the different types of messages to guide conversation as described in the activity Diagram presented in Figure 35, and refining the acceptance region of the description by further questioning the person participating in the dialogue until a small enough region is achieved.

Bibliography

[1] Christopher Alexander. *The nature of order: an essay on the art of building and the nature of the universe. Book 4, The luminous ground.* Center for Environmental Structure, 2002.

[2] Brandon Bennett. *Spatial Vagueness*, pages 15–47. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[3] Luka Bradeško and Dunja Mladenić. A survey of chatbot systems through a loebner prize competition. In *Proceedings of Slovenian language technologies society eighth conference of language technologies*, pages 34–37. Institut Jožef Stefan Ljubljana, Slovenia, 2012.

[4] Bénédicte Bucher, Gilles Falquet, E Clementini, and M Sester. Towards a typology of spatial relations and properties for urban applications. *Usage, Usability, and Utility of 3D City Models–European COST Action TU0801*, page 02010, 2012.

[5] Laura A Carlson and Eric S Covey. How far is near? inferring distance from spatial descriptions. *Language and Cognitive Processes*, 20(5):617–631, 2005.

[6] Christy Cheung, Matthew Lee, and Xiaoling Jin. Customer engagement in an online social platform: A conceptual model and scale development. 2011.

[7] Eliseo Clementini. Directional relations and frames of reference. *GeoInformatica*, 17(2):235–255, 2013.

[8] Eliseo Clementini and Giampaolo Bellizzi. A geospatial application framework for directional relations. *ISPRS International Journal of Geo-Information*, 8(1):33, 2019.

[9] Eliseo Clementini and Roland Billen. Modeling and computing ternary projective relations between regions. *IEEE Transactions on Knowledge and Data Engineering*, 18(6):799–814, 2006.

[10] Eliseo Clementini, Jayant Sharma, and Max J Egenhofer. Modelling topological spatial relations: Strategies for query processing. *Computers and Graphics*, 18(6):815–822, 1994.

[11] Anthony G Cohn and Jochen Renz. Qualitative spatial representation and reasoning. *Foundations of Artificial Intelligence*, 3:551–596, 2008.

[12] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[13] Richard Csaky. Deep learning based chatbot models. *arXiv preprint arXiv:1908.08835*, 2019.

[14] Andrea F Daniele, Mohit Bansal, and Matthew R Walter. Natural language generation in the context of providing indoor route instructions. In *Proceedings Robotics: Science and Systems Workshop on Model Learning for Human-Robot Communication*, 2016.

[15] Bradley Efron and Robert J Tibshirani. *An introduction to the bootstrap*. CRC press, 1994.

[16] Max J Egenhofer and Robert D Franzosa. Point-set topological spatial relations. *International Journal of Geographical Information System*, 5(2):161–174, 1991.

[17] BSc FHNW Andrew J Frei. *Integrating a chatbot with a GIS-MCDM system*. PhD thesis, University of Salzburg, 2018.

[18] Lucas Freitas and Cláudio Campelo. Human spatial reasoning in everyday language: Inferring regions that describe spatial relations. 12 2020.

[19] Agnès Gryl, Bernard Moulin, and Driss Kettani. A conceptual model for representing verbal expressions used in route descriptions. In *Spatial language*, pages 19–42. Springer, 2002.

[20] Ehsan Hamzei, Stephan Winter, and Martin Tomko. Initial analysis of simple where-questions and human-generated answers. 2019.

[21] Christoph Hölscher, Thora Tenbrink, and Jan M Wiener. Would you follow your own route description? cognitive strategies in urban route planning. *Cognition*, 121(2):228–247, 2011.

[22] Zhe Hu, Jia Pan, Tingxiang Fan, Ruigang Yang, and Dinesh Manocha. Safe navigation with human instructions in complex scenes. *IEEE Robotics and Automation Letters*, 4(2):753–760, 2019.

[23] Jiyou Jia. Csiec: A computer assisted english learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems*, 22(4):249–255, 2009.

[24] Arbaz Khan, Maria Vasardani, and Stephan Winter. Extracting spatial information from place descriptions. In *COMP@ SIGSPATIAL*, page 62, 2013.

[25] Parisa Kordjamshidi, Martijn Van Otterlo, and Marie-Francine Moens. Spatial role labeling: Towards extraction of spatial relations from natural language. *ACM Transactions on Speech and Language Processing (TSLP)*, 8(3):1–36, 2011.

[26] John Lafferty, Andrew McCallum, and Fernando CN Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. 2001.

[27] David M Mark. A conceptual model for vehicle navigation systems. In *Conference Record of papers presented at the First Vehicle Navigation and Information Systems Conference (VNIS'89)*, pages 448–453. IEEE, 1989.

[28] David M Mark and Max J Egenhofer. Modeling spatial relations between lines and regions: combining formal mathematical models and human subjects testing. *Cartography and geographic information systems*, 21(4):195–212, 1994.

[29] Cynthia Matuszek, Evan Herbst, Luke Zettlemoyer, and Dieter Fox. Learning to parse natural language commands to a robot control system. In *Experimental robotics*, pages 403–415. Springer, 2013.

[30] Daniel R Montello. A conceptual model of the cognitive processing of environmental distance information. In *International conference on spatial information theory*, pages 1–17. Springer, 2009.

[31] Gilvan R. Neto. Como pessoas descrevem lugares?, 2019. Trabalho de Conclusão de Curso (Bacharel em Ciência da Computação), Universidade Federal de Campina Grande (UFCG), Paraíba, Brazil.

[32] Tri Nguyen, Mir Rosenberg, Xia Song, Jianfeng Gao, Saurabh Tiwary, Rangan Majumder, and Li Deng. Ms marco: a human-generated machine reading comprehension dataset. 2016.

[33] Kyo-Joong Oh, Dongkun Lee, Byungsoo Ko, and Ho-Jin Choi. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In *2017 18th IEEE International Conference on Mobile Data Management (MDM)*, pages 371–375. IEEE, 2017.

[34] California Commission on Peace Officer Standards Training and Michael Josephson. Becoming an exemplary public safety dispatcher, 2015.

[35] Sirous Panahi, Jason Watson, and Helen Partridge. Social media and tacit knowledge sharing: Developing a conceptual model. *World academy of science, engineering and technology*, (64):1095–1102, 2012.

[36] Kiran Ramesh, Surya Ravishankaran, Abhishek Joshi, and K Chandrasekaran. A survey of design techniques for conversational agents. In *International conference on information, communication and computing technology*, pages 336–350. Springer, 2017.

[37] Gudula Retz-Schmidt. Various views on spatial prepositions. *AI magazine*, 9(2):95–95, 1988.

[38] Daniela Richter, Stephan Winter, Kai-Florian Richter, and Lesley Stirling. How people describe their place: Identifying predominant types of place descriptions. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information*, pages 30–37. ACM, 2012.

[39] Bodo B Schlegelmilch and Tina Claudia Chini. Knowledge transfer between marketing functions in multinational companies: a conceptual model. *International Business Review*, 12(2):215–232, 2003.

[40] Anthony Stefanidis, Andrew Crooks, and Jacek Radzikowski. Harvesting ambient geospatial information from social media feeds. *GeoJournal*, 78(2):319–338, 2013.

[41] Ilya Sutskever, Oriol Vinyals, and Quoc Le. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, 4, 09 2014.

[42] Ben Talbot, Obadiah Lam, Ruth Schulz, Feras Dayoub, Ben Upcroft, and Gordon Wyeth. Find my office: Navigating real space from semantic descriptions. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5782–5787. IEEE, 2016.

[43] Stefanie Tellex, Nakul Gopalan, Hadas Kress-Gazit, and Cynthia Matuszek. Robots that use language. *Annual Review of Control, Robotics, and Autonomous Systems*, 3:25–55, 2020.

[44] Martin Tomko and Stephan Winter. Pragmatic construction of destination descriptions for urban environments. *Spatial Cognition & Computation*, 9(1):1–29, 2009.

[45] Meng-Han Tsai, James Yichu Chen, and Shih-Chung Kang. Ask diana: A keyword-based chatbot system for water-related disaster management. *Water*, 11(2), 2019.

[46] Alan M Turing. Computing machinery and intelligence. In *Parsing the Turing Test*, pages 23–65. Springer, 2009.

[47] Maria Vasardani, Martin Tomko, and Stephan Winter. The cognitive aspect of place properties. In *International Conference on GIScience Short Paper Proceedings*, volume 1, 2016.

[48] Maria Vasardani and Stephan Winter. Place properties. *Advancing geographic information science: The past and next twenty years*, pages 243–254, 2016.

[49] Richard Wallace. The elements of aiml style. *Alice AI Foundation*, 139, 2003.

[50] Richard S. Wallace. *The Anatomy of A.L.I.C.E.*, pages 181–210. Springer Netherlands, Dordrecht, 2009.

[51] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, 1966.

[52] Stephan Winter and Yunhui Wu. Towards a conceptual model of talking to a route planner. In *International Symposium on Web and Wireless Geographical Information Systems*, pages 107–123. Springer, 2008.

[53] Wlodek Zadrozny, Malgorzata Budzikowska, J Chai, Nanda Kambhatla, Sylvie Levesque, and Nicolas Nicolov. Natural language dialogue for personalized interaction. *Communications of the ACM*, 43(8):116–120, 2000.

[54] Changqing Zhou, Pamela Ludford, Dan Frankowski, and Loren Terveen. *Talking about place: An experiment in how people describe places*. na, 2005.