Nearness as Context-Dependent Expression: An Integrative Review of Modeling, Measurement and Contextual Properties

Marc Novel¹, Rolf Grütter¹, Harold Boley², and Abraham
Bernstein³

¹Swiss Federal Institute for Forest, Snow and Landscape Research
WSL

²University of New Brunswick, Faculty of Computer Science

³University of Zürich, Department of Informatics

March 31, 2020

Nearness expressions such as “near” are context-dependent spatial relations and are subject to the context variability effect. Depending on the provided context, “near” has a different semantic extension. We perform a literature review to identify the effect of context on “near”. To integrate the insights from different disciplines, we apply Turney’s contextualisation framework which distinguishes between two types of features: primary and contextual. Primary features are the qualitative and quantitative distance measures and contextual features are the context factors used

to determine a threshold on the nearness measurements. Additionally, we identify the appropriate features for different spatial tasks discussed in the literature. By doing so, we seek to build a foundation for a context-dependent model for “near”.

Contents

1	Introduction	3
2	Context and “near”	7
2.1	Incomplete Information and Contextualized Interpretation of “near”	8
2.2	Contextualized Nearness Models	10
3	Methodology of our Review	16
4	The Spatial Concept of “near”	20
4.1	The Locative “near”	20
4.2	Properties of “Near”	23
4.3	Semantics of “near”	25
4.3.1	The Ambiguity of “near”	28
4.3.2	The Frame of Reference for “near”	33
4.4	Functional Relations for “near”	35
5	Contextual Features of a Model of “near”	37
5.1	Topological Properties of the Domain	38
5.2	Primary Features	39
5.2.1	Qualitative Approaches to “near”	39
5.2.2	Quantitative Approaches to “near”	42
5.3	Granularity	47
5.4	Context Factors	50
5.4.1	Associating CFs only with the LO	51
5.4.2	Associating CFs only with the RO	52
5.4.3	Associating CFs with both Objects (LO & RO)	54
6	Contextualized Tasks for “near”	59
7	Conclusion and Outlook on Future Research	61

1 Introduction

The spatial concept of nearness, or proximity, is used to identify either a spatial relationship of nearness between two entities such as *near(market, London Bridge)* or an object and its relative location in space (*market(near, London Bridge)*). Probably all languages have a way to communicate this spatial relationship and in English, the most common way is by using the expression “near”.

Depending on the research field, the expression “near” is identified as a qualitative spatial expression (geography and qualitative spatial reasoning (Wallgrün, Klippel, & Baldwin, 2014))

or as a locative (linguistics (Herskovits, 1986)).

The expression “near” is used to describe spatial proximity in narrative text (route or place descriptions), geographic web queries (Sanderson & Kohler, 2004), or robot instructions. Consider the following examples:

- In a **query**, “near” plays an important role when asking for information about a location, a place, or a region: *Football stadiums near Stratford train station, Hotels near Heathrow Airport, Indian restaurants near Covent Garden*. This is very common for geographic queries on the web.
- The expression “near” plays a prominent role for **place descriptions**, such as “*at Birran-rung Marr near Federate Square*”, “*near corner West fourth St. and Prime Avenue*”, or “*The village is to the south of the Ski hill, near the highway exit*”, where it relates two spatial objects. Additionally, “near” connects already known information with new information to provide an informed place description.
- The role of “near” in **route descriptions** such as “*When near the church, turn left and follow Bank Street until you are near the train station*” is to link the spatial objects together for cognitive adequacy (Gryl et al., 2002).
- Spatial expressions are used to **instruct robots** in human-robot interaction. For instance, “*get the knife near the plate*”, “*go to the statue near the market*” (Crangle & Suppes, 1989; Kuipers, 2000).

If the goal is to parse these kinds of sentences, to extract the data, or generate sentences based on spatial data, a formal model is needed. Such a model establishes a semantic connection between an expression denoting the spatial relation of *nearness* and a precise denotation.

We follow the research done within the field of geographical information science (GIScience), which assumes a direct relation between the interpretation of “near” (i.e. characteristics of the region, the distance between the object to be located and the reference object) and context (Brennan & Martin, 2012; Hawes, Klenk, Lockwood, Horn, & Kelleher, 2012; Wallgrün et al., 2014; Xu & Klippel, 2013).

For instance, in a densely populated area, such as London or Manhattan, a nearby petrol station is expected to be within a distance of 5 min travel time, whereas in a rural area, a petrol station within a distance of 30 min travel time may still be considered to be near. In the case of

“near for a daily commute”, the situation is reversed. A daily commute with a travel time of 30 min might be considered as “near” in an urban context, but not “near” in a rural one. In the following examples “|” means “given that”:

- *travelling 30 min is near* = false | Context = urban, petrol station, car
- *travelling 30 min is near* = true | Context = rural, petrol station, car
- *travelling 30 min is near* = true | Context = urban, daily commute, public transport
- *travelling 30 min is near* = false | Context = rural, daily commute, car

This phenomenon that different interpretations of “near” can be used by different subjects in different contexts is called the *context variability effect* and has the following consequences:

- In case “near” is interpreted in terms of a distance measurement, then the context-variability effect is observed. Thus a nearness model needs to allow for different interpretations in different contexts.
- Not vagueness but context dependency is the key element for understanding “near” and a nearness model needs to take context variability into account. A formal model for “near”, which takes contextual information into account, is expected to have a better predictive power than a model without contextual information (Gurevych, Porzel, & Malaka, 2006).

Historically, the importance of context was not always universally accepted. Some approaches, such as qualitative spatial reasoning (QSR), usually emphasize geometric or topological properties of near, whilst treating context with lower importance (Frank, 2006). We identify at least three reasons for the negligence of context: (i) There is neither a formal model of context nor a universal definition or a characterization of how context is constituted¹. (ii) Very often it is assumed that a uniform interpretation of “near” is possible. (iii) The phenomenon of vagueness is sometimes considered to be more fundamental than context dependency and subsequently, context is largely ignored.

Nevertheless, the context-sensitive nature of “near” has now been investigated for a while both in cognitive science and in linguistics. In cognitive science, the theory of cognitive maps studies how geographic objects can have different extensions, based on their *Gestalt* or mathematical properties and the contextual situation. Furthermore, research in cognitive maps tries to

¹See Bazire, Bazire, Brézillon, and Brézillon (2005) where 150 different definitions of context are examined without being able to provide a unified and consistent definition.

determine the constraint factors which distort or transform the distance estimation for “near” in a given situation (B. Tversky, 2003). In linguistics, starting with Herskovits (1986), the context-dependency of “near” has been widely discussed (for formal semantics see (Zwarts, 2017), for spatial language research see (Burigo & Coventry, 2010; Burigo, Coventry, & Cangelosi, 2016; Coventry, 2013; Coventry & Garrod, 2004)).

In the field of geographical information science (GIScience), a systematic relation between the interpretation of “near” and the context has been recognized only recently (Brennan & Martin, 2012; Hahn, Fogliaroni, Frank, & Navratil, 2016; Hawes et al., 2012; Minock & Mollevik, 2013; Skoumas, Pfoser, & Kyrillidis, 2014; Wallgrün et al., 2014; Xu & Klippel, 2013). Whereas especially Brennan and Martin (2012) recognized that “near” not only depends on a distance. Depending on the context, other measurements are sometimes more relevant than distance.

Hence we recognize that, for the interpretation of nearness, a direct mapping to a distance metric is cognitively not adequate. To alleviate this problem, we introduce a way of formulating a context-dependent model for “near”. To do so, we use Turney’s contextualisation framework (Turney, 1996b), where two dimensions are used for context-dependent modeling.

In case of “near”, the first dimension is identified with different distance measurements. These measurements have the property of being able to evaluate if something is near or not. However, the prediction accuracy of using distance measurements is limited and therefore cognitively not adequate.

The second dimension is the contextual information. Though context does not provide any metrics for evaluating whether something is near or not, it provides relevant information for the determination of what is near.

Based on these two dimensions, we provide an integrative and systematic review (Whittemore & Knafl, 2005) about the role of the context for “near”.

The remainder of the paper is structured as follows: In section 2 we give an overview of the different contextualized models for “near” suggested in the literature. We identify that a context-dependent model for “near” usually contains two elements: a *primary* and a *contextual* element. Furthermore, we discuss the notion of vagueness, and discuss how to separate vagueness from context-dependency. This section informs us about the methodology of this review, which we discuss in section 3.

In section 4 we describe the concept of “nearness” and introduce the English locative “near” and identify its properties. These properties, in turn, show which requirements a model needs to meet, in order to be considered to be cognitively adequate. We also describe how “near” is ambiguous, as “near” can be used for *spatial descriptions* or *object identification* (spatial reference). We discuss a denotational semantics for the two interpretations of “near”: “near as a region” and “near as a distance”. Additionally, we discuss the pragmatic effect of functional relations.

In section 5, we present the features of a contextualized model for “near”. For this we conduct an integrative literature review and classify the features found into primary and contextual features. Primary features are different qualitative or quantitative distance measures, represented as scales. Subsequently, we identify and enumerate the context factors (contextual features of the nearness model). Context factors (CFs) are information from the context and we describe them as a transformation over primary features. We also encode how the different context factors transform a region or a distance.

In section 6 we discuss how different spatial tasks and the features are related to each other in terms of functional relations.

The paper concludes with section 7, where we summarize our findings and make suggestions for further research.

2 Context and “near”

According to Dey (2001), context is defined as “*any information that can be used to characterize the situation of an entity*”, where *situation* means *circumstances* or *conditions*. To provide a context-sensitive model for “near”, the main task is to identify and determine the contextual conditions which, in turn, determine the interpretation of “near”. For instance, the interpretation of the assertion “*there is a gas station nearby*” has a different interpretation when uttered in a city than when uttered in a rural area, since there are fewer gas station in a rural area. In this case, the situation of “near” can be made explicit by relating it to the population density and the distance measure of travel-time by car. Hence, a context-sensitive model for “near” qualifies if the model provides extra information to specify the conditions or circumstances for which “near” applies. Such specification could, for instance, distinguish “*near for a day-trip*” from “*near*”

for a weekend-trip". In these cases, depending on the conditions, "near" is interpreted differently. By providing supplementary information and reducing uncertainty, a contextualized model is expected to make better predictions than a model that disregards context. Hence, instead of treating "near" as a vague expression, we treat "near" as a context-dependent expression and in section 2.1 we discuss the implications of doing so.

In section 2.2 we give an overview of the different approaches to modeling "near" as a context-dependent expression adopted in psychology, cognitive science, linguistics, and geographic information science. The insights gained from this overview inform the methodology used in our interdisciplinary review, discussed in section 3.

2.1 Incomplete Information and Contextualized Interpretation of "near"

Traditionally, the main source of difficulty when defining a formal model for "near" has been identified in the *vague nature* of the expression "near" (for an overview of vagueness and "near" see Vasardani, Winter, and Richter (2013)). The *vague nature* is identified with the absence of a sharp boundary between "near" and "not near".

One way to deal with this kind of problem is by discretising "near" into grades of nearness via fuzzy logic or fuzzy sets. Research then is conducted to provide for an empirically adequate discretisation function to answer the following problems: *The problem of vagueness*: An expression is considered to be vague, when it is susceptible to the Sorites paradox. For "near" the Sorites paradox can be constructed (P. Fisher, 2000): if we are near location *A* and move one meter away, then we are still near the location *A*. If we move another meter further away, again, we are still near. If we repeat this procedure, we always consider to be near the location *A*, even though, in fact, we are no longer near the location. Thus, drawing a line between near and not near is problematic. The *problem of ambiguity*: "near" has many different meanings: "near for a day ride", "near for having a drink" and the *problem of borderline contradictions*: objects can simultaneously be "near" and "not near" at the same time.

In the case of context-sensitive models, the view on the problem of *vagueness* is shifted. Instead of dealing directly with these problems, a context-dependent model determines a threshold with an appropriate distance measure (or scale) and the information from the context. Since

the nearness relation is *tolerant*, it is insensitive to small changes, which is then perceived as *vague* (Denofsky (1976) calls it *vagueness band*). *Tolerance* is determined by the context (granularity of the measurements and contextual information) and can shift. Research is conducted to empirically determine the appropriate tolerance with respect to the context. Comparing a fuzzy-modeling to a contextualized-modeling approach looks as follows:

The Problem of Vagueness: In the case of fuzzy logic, the expression “near” is first crisped and then a contextualized interpretation is provided (Barouni & Moulin, 2015; Gahegan, 1995). A contextualized model, instead, will suggest that in absence of contextual information, an expression containing “near” is underspecified and “near” cannot be fully interpreted (Gurevych et al., 2006).

By adding contextual information, the actual interpretation of “near” becomes more defined. Examples are “near for a day ride”, “near for grocery shopping”, “near for a sparsely populated area” (Denofsky, 1976; Minock & Mollevik, 2013).

The Problem of Ambiguity: In fuzzy logic vagueness and ambiguity are treated as the same problem. A context-dependent model treats vagueness and ambiguity differently. In absence of a context, multiple interpretations of “near” are possible.

Consequently, contextual information eliminates possible interpretations, and relevant contextual factors help to disambiguate “near” (Cai, 2007; Gurevych et al., 2006).

The Problem of Borderline Contradictions: Fuzzy logic tries to discretise “near” and assign a degree of nearness. A context-dependent model solves this problem differently. In absence of a context, an object can be “near” and “not near” at the same time. This usually occurs when no clear line can be drawn between “near” and “not near”, nor between the opposite pair “near” and “far” with respect to a context. Hence, by providing context, the contradiction can be resolved: “near” in context C_1 and “not near” in context C_2 (Cai, 2007; Minock & Mollevik, 2013).

By reframing vagueness issues into context-dependency issues, we expect to address all the problems a vague-model will address. As mentioned earlier, we also assume that a context-dependent model is expected to have a better predictive power as it reduces uncertainty by taking additional contextual information into account. The uncertainty reduction effect has been verified in a large body of psychophysical studies, where it has been shown that the judgment

accuracy of respondents increases with more available information (Wagner, 2006).

2.2 Contextualized Nearness Models

To model the contextual variability of “near”, there are a few approaches found across disciplines. This section provides an overview of these different approaches from the literature and we will summarize the common features.

The overview of the different contextualized nearness models will be the starting point for our problem identification process. Based on this overview we will choose an appropriate methodology for the review, which we discuss in the following section (section 3).

The following approaches can model the context dependency of “near”: statistical, cognitive, logical (ontological), and (neuro-) fuzzy. Their common goal is to capture the element of context variability (i.e. “near” having a different interpretation in varying contexts)

In the computer science & GIS-science literature a distance and a contextual dependent threshold on that distance are used to model the context dependency of “near”. An early example of calculating thresholds for “near” is Denofsky (1976), where the calculation of the distance for “near” is based on different factors such as *purpose*, *(largest) size of the spatial objects*, *(maximal) absolute distance*, *relative distance* and *standards* (distance compared to the adjacent objects). To calculate a threshold, Denofsky calculates a “near” threshold for each factor individually and combines each of them via geometric averaging to a global threshold. The individual thresholds themselves are represented as distributions and therefore facilitate the association of a probability for a global threshold. Such as for distance X with the factors $Y, Z, \dots$ there is a 70% chance that the observer will classify this resulting distance as “near”.

The approach of determining a contextually defined threshold for “near” has been further developed within the field of GIS-science. Instead of using *absolute distance*, Worboys (1996) uses relative distance to measure the distance for “near”. He defines the relative distance between two spatial objects A and B as the Euclidean distance between A and B , divided by the mean of distance between A and every other object in consideration. Gahegan (1995) also uses relative distance, but not exclusively so. Besides the relative distance he uses absolute distance and the combination of absolute and relative distance in conjunction with the following context factors:

scale, density of spatial objects, attractivity, and reachability.

Neither Denofsky (1976), nor Worboys (1996) used any experimental data to validate their approaches, while Gahegan (1995) conducted a limited psychometric test to validate his approach.

Within the field of GIS-science, there are also approaches where distance is not considered to be relevant for “near”. Instead of relying on a distance, “near” is represented as a region. That is, instead of calculating a distance threshold, a cognitive region is defined as being applicable for “near” and can be determined qualitatively (see the details below).

Another alternative to model “near” within GIS-science is suggested by Yao and Thill (2005), where the distance is represented as a 5-point ordinal scale (“*very near*”, “*near*”, “*normal*”, “*far*”, “*very far*”). The model used to predict “near” is an ordered logit regression (OLR) model, with the dependent variable being the ordinal scale of nearness, where the independent variables present the context factors and distance. Additionally, Yao and Thill distinguish between *objective* and *subjective factors*. *Objective factors* are independent of the person perceiving the distance, such as *scale of the study area, type of activity, reachability, intervening opportunity, transportation mode, traffic condition* and are geometrically modeled. *Subjective factors*, on the other hand, such as *familiarity with the area and the route, the perceiver’s time and financial budget, the perceiver’s demographic characteristics, socioeconomical characteristics*, are modeled extra-geometrically.

Within spatial language research, where the path descriptions of the image schema of the figure-ground relation in space is investigated, it was noted early on that spatial relations are not solely determined by geometrical factors, but extra-geometric or functional factors play also an important role in determining “near”.

Vandeloise (1991) argues that “near” is a distance and depends (mostly) on what he considers functional factors such as *ease of access, size, mode of transport, ability of agent* and physical factors, such as *topography, or accessibility*.

The notion of functional factors has been further investigated by Ferenz (2000). He conducts an experiment showing a picture of a room with a TV and different couches, where some couches are facing in the direction of the TV, while other couches face other directions. Respondents only considered the couches facing in the direction towards the TV to be near. Couches not facing in the direction of the TV were not considered to be “near”, even though the couches were closer to the TV. From this experiment Ferenz concludes that in order to determine “near”

between two objects, the functional notion of *the possibility of them to interact with each other* is more important than the geometrical notion of distance.

This finding has been further corroborated within Coventry and Garrod's functional geometric framework (Coventry & Garrod, 2004). The functional relational framework states that locatives, in general, and "near", in particular, depend on geometrical and functional properties. However, each role needs to be determined empirically. In their research they found the following geometrical context factors as relevant for "near": relative size of the objects determines the distance for "near" (*a pencil near an eraser* and *a car near a garage*), or movability (*the bicycle is near the mailbox* not *the mailbox is near the bicycle*), and purpose (*the ball is near the hole* in the case of a game of snooker is different from a golf course). This has also been experimentally verified by Carlson and Covey (2005) through experiments mostly conducted within figural-space.

Within psycholinguistic research *spatial templates* have been used to represent graphically on a template the acceptability associated with "near". The spatial template is centered around the reference object (RO, see section 4.1 for a definition) and identifies the nearness acceptability for each point or quadrant within the spatial template. Logan and Sadler (1996) generated experimental data by asking subjects how well each cell fits with the linguistic expression "X is near the RO". They noticed that the acceptability decreases the further away the X is placed from the RO. However, the acceptability rating again increased at the edges, which leads to the conclusion that not only distance might play a role for determining "near".

In psychology, distance estimation and distance perception has been studied extensively. Wagner (2006) reports of numerous studies investigating how the different contextual factors such as *instructions, cue conditions, memory vs. direct judgment, the range of stimuli, or judgment method* influence the distance estimation of respondents. To calculate the distance estimation Wagner uses magnitude estimation, where the scaling constant and exponent of the magnitude estimation are now dynamic and more like functions of judgment conditions. While these studies give an insight into how the different context factors have an influence on the subjective distance perception, they do not answer the question of how to determine a threshold for "near".

Multidimensional scaling (MDS) is another statistical method used in psychology to extract

dimensionality from relative similarities in a multidimensional space. In scalar MDS, scaling can be achieved via Minkowski distances. The different contexts form an n -dimensional space, which is then scaled down to a 2-dimensional space (Hourihan & Jones, 1979). For non-scalar MDS only the order of similarities is relevant and the fitting is measured in terms of STRESS values (STRESS measures the difference between the observed values and the estimated values).

Experiments have shown that images generated from non-metric representations require little fitting (STRESS) for “near” (McNamara, Hardy, & Hirtle, 1989). In an n -dimensional space the Euclidean axiom of the triangle inequality does not hold anymore. Since the MDS representation forms an n -dimensional space and “near” does not fulfill the axiom of triangle inequality, a quantitative interpretation of “near” is possible (Wagner, 1985). However, the property of asymmetry is problematic for multidimensional scaling (Hourihan & Jones, 1979). For this reason transformations are applied to achieve asymmetry (McNamara & Diwadkar, 1997).

Pragmatic theories exploit the fact that knowledge is usually never isolated but depends on knowledge previously acquired. This can be previous utterances (such as a dialog, or interaction with the environment), or the current task or goal of the agent. Since the interpretation of “near” depends on this kind of information, “near” for one agent is not necessarily the same as for another agent. By taking additional information into account contextual knowledge can be made available by reconstructing the communication process via additional information. Pragmatic theories model how a priori information plays a role when interpreting context-dependent expressions. Studies reconstructing the a priori information with “near” in order to interpret the information with respect to the context are either based on *information theory*, *Bayesian reasoning*, or *neuro-fuzzy techniques*.

Worboys (2003) uses channels from *information theory* (Barwise & Seligman, 1997; Shannon, 1948; Shannon & Weaver, 1949) to model such information. In his case “near” is defined as a neighborhood region with the actual extension based on the notion of nearness in terms of the agent’s belief system. Each such neighborhood region then acts as a classifier for “near” and has its own information channel. Since the classifier depends on an agent, each classifier will produce a different output for each agent. An additional channel is set up which mediates between the different outputs of each information channel.

With the *Bayesian approach* developed by Ma, Kalashnikov, and Mehrotra (2008), “near” is

modeled as an uncertain event location described by human writers and represented as a probability density function (PDF). The goal is to determine $f(x, y|s, t)$, representing the PDF of $f(x, y)$, with x, y being the values of an uncertain location in $\mathbb{R}^2$, given an expression s with the locative “near” and an event t . As an example, the PDF of the statement “A traffic accident near the World Trade Center” is represented as $f(x, y|s = \text{near}(\text{WTC}), t = \text{“traffic accident”})$ ². To determine the PDF, first the locative “near” is mapped to a (truncated) Gaussian density distribution with its center at the RO. The density distribution serves as a Bayesian prior and the updating is performed by conditionalising with respect to the event t .

In the case of *neuro-fuzzy techniques*, fuzzy membership rules are learned with respect to context. In Barouni and Moulin (2015) a neural net is defined in terms of linguistic variables as its input layer and class memberships as its output layer. For instance, *Euclidean distance* is assigned to one of the sets *short, medium, far*, the variable *traffic* is assigned to *light, medium, heavy*, and the variable *familiarity with the site* is assigned to *weak, average, good*. These sets then are fuzzified according to a learned rule and assigned a class membership of the following categories *very short, short, medium, far*.

In the case of the *question answering* procedure: the context is reconstructed by asking the user for contextual information, instead of trying to reconstruct the previous context via some learned rules.

Systems asking the user for contextual information usually implement these via question answering systems (Q/A). These systems explicitly request feedback from the user to determine the contextual conditions on which “near” is used. For instance Minock and Mollevik (2013) use Q/A to learn distance thresholds for different contexts. By doing so, appropriate recommendations for restaurants or touristic places with respect to the context can be provided.

In a series of papers Cai and colleagues (Cai, 2007, 2009; Cai, Wang, & MacEachren, 2003; Cai, Wang, MacEachren, & Fuhrmann, 2005) combine Q/A with planning. As they made the observation that “near” depends on the goal of an agent, they query via Q/A the task and goal of the user and adapt geographical information system (GIS) accordingly to the task. Finding the goal of the agent is then treated as a planning problem based on the information provided

²The Port Authority Police Department (PAPD) reports from the 11 September 2001, is used as natural-language corpus.

by the adapted GIS.

Another approach to model “near” is in terms of qualitative spatial regions. For instance, Grütter, Scharrenbach, and Waldvogel (2010) determine a region for “near” by partitioning an area into regions with specific qualitative properties. The result of such a partitioning is a mereology (or a mereotopology), which models “near” based on qualitative properties. Qualitative properties, which have been investigated so far, are services rendered at different levels of administrative organizations (Grütter et al., 2010), the uniform behavior of commuters (Grütter, Helming, Speich, & Bernstein, 2011).

Another way to model cognitive regions is that of Brennan and Martin (2012), who introduce the notion of *impact area*. An impact area is based on *influence area* (Kettani & Moulin, 1999) and is defined as the part of the space surrounding an object. Euclidean geometry is used to determine an influence area of an object and the influence area has the same shape as the object itself. An impact area is an influence area with added context information (i.e. a mapping from influence area and context to impact area). The context itself is determined via a domain expert, who decides which contextual information is relevant and collects that information for the specific task.

Another way is to represent functional relation is via map layers (Hahn et al., 2016). Depending on the task at hand a different layer is needed. In case the task is to find a pub nearby, then a different layer is shown than when the task is hiking, where the map layer with hiking paths is needed. Usually the right layer is chosen by the agent performing a task

In the domain of *ontologies*, the role of context needs to be made explicit in order to gain adequate spatial knowledge representation. There are at least two strategies to model the context-dependency. One is a “two-level” approach developed by Bateman and colleagues (Bateman, 2010a, 2010b; Bateman, Hois, Ross, & Tenbrink, 2010), where an abstract level and a concrete level is introduced. These two levels are realized via modularized ontologies (for the abstract level an individual ontology and also for each context an own ontology). The information between different ontologies is mediated via the ϵ -connection calculus and the two levels are specified as follows:

- Abstract Level provides underspecified qualitative (mereotopological) descriptions to accommodate the meaning across all possible contexts. On this level, only functional rela-

tions, introduced by the spatial preposition, are relevant. For instance “near” and “far” are described as opposite spatial relations.

- Concrete Level provides the primitives so that the abstract functional relations can be contextualized for a specific situation. For instance, if precise distance measures are needed for the “near”/“far” distinction, this is realized on this level. In the case “near” should be interpreted as a region or a distance, this is also interpreted on this level.

Another approach also using ontologies is that of Pareti and Klein (2011). Techniques of ontology evolution and ontology alignment are used to reconstruct the communication process. The nearness thresholds are learned from data.

3 Methodology of our Review

As seen in the previous section, there is a great variety of approaches to modeling the context-dependent expression “near”. We mainly attribute this diversity to the interdisciplinary nature and the different aspects associated with the concept of “near”.

With respect to the different nearness models, we identify models based on theoretical and empirical sources, where the empirical models use either experimental or non-experimental methodologies. Experimental data is gathered by having the tested persons judge how likely they will use a nearness expression for a given situation, whereas non-experimental data is extracted from NL descriptions, such as travel descriptions from internet blogs.

We also identify a large variety of the different features in these models. Some treat “near” as a distance, which can either be a relative distance or on a metrical scale such as Euclidean distance. Other models do not treat “near” as a distance but as a region, which is defined in topological or mereotopological terms. Depending on how “near” is modeled, this has further implications. For instance, some studies found that direction of the spatial objects has an influence on the interpretation of “near”, but if “near” is modeled as a region, the information about direction cannot be represented.

With respect to the information from the context, we see a wide variability of different CFs investigated (for a definition of CF see section 5.4). To the best of our knowledge, there is no consensus across the literature as to the definition how a minimal model for “near” is consti-

tuted.

Another dissimilarity among the models is the variety of the spaces taken into consideration. We adopt the distinction proposed by Montello (1993), where he partitions the space into figural, vista, environmental, and geographical space with the following properties:

- **Figural space** is smaller than the (human) body and can be observed directly from one place, such as *pictures, small objects*.
- **Vista space** is larger than the body and can visually be understood from a single place, such as *single rooms, town squares, small valleys and horizons*.
- **Environmental space** is larger than the body or the surroundings and cannot be understood without moving around, such as *buildings, neighborhoods and cities*.
- **Geographical space** is much larger than the body and cannot be experienced directly through locomotion but is learned via symbolic representation, such as *states, countries or the solar system*.

While most theoretical and non-experimental approaches are primarily concerned with environmental and geographical spaces, experimental approaches, on the other hand, mostly investigate figural spaces and it is not exactly clear how the results from figural space can be projected into larger space (Sharma, Flewelling, & Egenhofer, 1994). Sometimes it is even argued that the empirical results from figural space are not applicable to larger space (van der Zee, Adams, & Niemi, 2009).

But despite all the different ways of how to model the context-dependency of “near”, all models try to answer the following two questions:

- Which elements have an influence on the interpretation of “near”?
- How do these elements influence the interpretation of “near”?

We conducted a literature review to provide answers to these questions. The goal of this literature review is to introduce the possible features for a nearness model which determines what conditions “near” relies upon to make an adequate prediction. In order to determine the criteria for such a model, the features of the different contextualized models extracted from the literature need to be made comparable and the conditions for “near” to be defined.

The methodology used for our literature review is an integrative review (Whittemore & Knafl, 2005) since we identify a large variety of examples in terms of methodological differences in

how to model the context-dependency of “near”. This integrative review allows us to integrate the literature with respect to the question about the role of context for “near” and relevant contextual information (see the two questions above).

To the best of our knowledge, a review is missing which integrates the various insights, found in the different contextual models across different disciplines with respect to these two questions. So far, most literature attempting an interdisciplinary review only addresses the first question. That is, they only compare the different context factors with respect to (relative) distance but not with respect to other distance measures or regions. B. Tversky (1992) classifies CFs with respect to *cognitive perspective*, *hierarchical organization*, and *cognitive reference point*. Within spatial language research the different CFs are compared (Carlson & Van Deman, 2004; Coventry & Garrod, 2004; Regier & Carlson, 2001; van der Zee et al., 2009). Yao and Thill (2005) compare different CFs by distinguishing between subjective and objective CFs. Wagner (2006) studies how different CFs influence distance judgments. Within the field of GIS-Science, Cai (2007) investigate the interaction between different CFs, and Hall and Jones (2012) investigate the correlation between distance measures and CFs.

To the best of our knowledge, only Brennan (2009) and Brennan and Martin (2012) compare different distance measures and CFs for “near”. They are mostly interested in the question of developing proximity measures where distance becomes a modifying factor among other factors. For this reason, they do not make a systematic comparison of different CFs and their effect on the various distance measures.

We collected papers which provide an answer to the aforementioned questions. We then broadened our search by looking for papers discussing at least one of these questions. As a search method, we performed a networking search. First, we performed a forward networking search with the initial pool of the papers discussed in the previous section. In that pool of papers, we looked for references which also discuss the concept of “near”. If so, we checked whether they provided an answer to at least one of our research questions. If this was the case, we added this paper to the pool. We reapplied this procedure for the newly discovered papers and repeated the procedure until no new papers were discovered. Additionally, on each search step, we also checked on *Google Scholar*³, to find out whether the papers selected in our pool

³google.scholar.com

were cited by a paper, addressing at least one of our research questions. We reapplied this procedure for the new papers and repeated it until no new papers were discovered.

The final sample of this integrative review includes empirical and theoretical reports including a variety of methods across disciplines. We extracted for each paper the discussed elements on “near” and how these elements have an influence on “near”. We took note if the discussed elements were validated empirically.

Since we are looking for “near” in large-scale environments, we excluded papers that explicitly exclude investigations on environmental or geographical spaces. We also excluded unpublished PhD theses and reports, unless they contained important insights. We considered unpublished works to be important in the case they have been cited more than once in the literature.

During our literature review it became apparent that the context-dependency of “near” depends on two elements: an element which can be considered in isolation and is independent of other features (distances, regions) and a context-variant element which is dependent on the independent feature. To provide for a synthesis of our integrative review we adopt Turney’s contextualisation framework (Turney, 1996b), which distinguishes between primary, contextual, and irrelevant features. Primary features are useful for nearness classification when considered in isolation without regard to the other features. Contextual features are not useful in isolation, but are useful when combined with other features. Irrelevant features are not useful for classification.

Primary features for “near” are the different kinds of distances and regions which are useful for classification tasks: “near” as within a certain distance threshold or “near” as within a certain region.

Following Turney (1996b) these primary features can be used for classification tasks, but will make better predictions when contextual features are also considered. These contextual features in turn, have an effect on the primary features. To account for this, we encode for each contextual feature the effect on the primary feature. Furthermore, we encode the different primary features with respect to their scale level, as this is their common denominator.

This allows the synthesis of the diverse methodologies and insights on contextualized nearness models into a unified representation of its features. Such a synthesis will serve as a basis

for a formal model of “near”.

We synthesize the result of our literature review by presenting a structured table of features as follows: the vertical dimension represents the primary features and the horizontal dimension represents the contextual features (see table 7, 8, 9, and 10 in the appendix).

4 The Spatial Concept of “near”

The linguistic expression “near” is a *locative* preposition (Herskovits, 1986) expressing nearness. Locatives are important in the spatial domain, where they are called *spatial concepts*. In section 4.1 we give an overview of the English locative “near” and in section 4.2 we describe the properties of the nearness relation which then serve as a basis for a cognitive adequate semantic model.

We then follow recent work in the natural language (NL) semantics of locatives by identifying two levels for the analysis of locatives: a semantic level and a functional level. In section 4.3, we discuss the semantic model, where we interpret “near” in a geometrical model (by means of region, distance, and direction). In this section we also discuss how “near” gives rise to an ambiguity as “near” can either denote a region $near_R$ or a distance $near_D$. In section 4.4 we discuss the non-geometrical but functional conditions for “near”.

The properties identified in this section are used to qualify the components of a model of “near” in section 5. In this section we also discuss the context-dependency of “near” and show how the context influences the denotation of “near” (context variability effect).

4.1 The Locative “near”

“Near” is a spatial preposition, called a locative (Herskovits, 1986), expressing the concept of *nearness* or *spatial proximity* and is the most common way to express spatial proximity in English, such as *Pubs near Big Ben*. In English, there are also other ways to express spatial proximity. It can be expressed by a synonym of “near”, such as by the following locatives: *close*, *close to*, *in the vicinity of*, *adjacent to*, *in the neighborhood of*.

In English, the expression “near” and some of its synonyms (most notably “close”) are sometimes ambiguous with respect to the assigned parts of speech (POS). “Near” can be a prepo-

sition, adjective, or adverb (see table 1). Additionally, the comparison form (positive, comparative, superlative) of the different POS can be formed, including the preposition (G. A. Miller & Johnson-Laird, 1976). The different POS become relevant, when we discuss the semantic properties of “near” in section 4.3.

POS	Comparison	“near”
Preposition	positive	“Pubs/the pub near Big Ben”
	comparative	“Pubs/the pub nearer Big Ben”
	superlative	“The pub nearest Big Ben”
Adjective	positive	“The road passes near Big Ben”
	comparative	“This road passes nearer Trafalgar Square than the other road”
	superlative	“This road passes the nearest”
Adverb	positive	“Come near to me”
	comparative	“Come nearer to me”
	superlative	“Come nearest to me”

Table 1: Comparative forms for “near”

Additionally *spatial proximity* can also be expressed by locatives such as “at”, “by”, “next to”, which express a more narrow *spatial proximity*, or locatives such as “behind”, “above”, which restrict the spatial proximity by introducing a direction: “beside” (near, but not above), “behind” (near, but not in front), “above” (near, but not beside or below) (Zwarts & Gärdenfors, 2015). Furthermore *spatial proximity* can also be expressed with directionals, such as “approach”, “reach”, “go toward”, or “go near” (Morrow & Clark, 1988), or demonstratives such as “this” or “here” (Kemmerer, 1999).

In this integrative review, we confine ourselves mostly to the English expression “near”, since it is the most frequently researched language. Especially for papers within the field of geography, qualitative spatial reasoning, psychology, and cognitive science, this is the case. Nevertheless, there are some cross-linguistic studies on near, where the English “near” is compared with its counterpart in French (Vandeloise, 1991; Vieu, 1991), Italian (Burigo & Coventry, 2010), German (Kracht, 2008), Finnish (van der Zee et al., 2009), Tsez, Muna (Austronesian), Saami (Kracht, 2008) or Yindjibarndi (Mark, Turk, & Stea, 2007). While many languages also encode “near” as a locative, other languages express *spatial proximity* via morphology. Also interesting is how languages make the distinction of “near” to the speaker, “near” to the hearer or “not near” to any speech participants (Kemmerer, 1999).

The locative “near” can have different signatures:

Two Arguments: The expression “near” usually takes two arguments. To get a spatial interpretation of “near”, the arguments are usually spatial objects. Thus “near” puts two spatial objects into a relation. The first argument is called the to-be-located-object (*LO*) and the second the reference object *RO* (Herskovits, 1986). In the sentence “*Pubs near Big Ben*”, “*Pubs*” is identified as the *LO* and “*Big Ben*” is identified as the *RO*.

To emphasize the fact that the spatial objects can be organized according to the figure-ground principle, sometimes the *LO* and *RO* are called *Figure* and *Ground* (Talmy, 2000). The same spatial object can serve as a figure or as a ground. For instance in the two sentences “*Pubs near Big Ben*” and “*Neighbours near pubs*”, the object “*Pubs*” serves as a figure in the first sentence and as a ground in the second.

In some cognitive studies the notions of *Trajector* and *Landmark* are used instead of *LO* and *RO*, referring to the “path” image schema by Johnson (1987) and Lakoff (1987).

One Argument: In some cases the *RO* is omitted in natural language as the *RO* implicitly refers to the location of the speaker: “*pubs nearby*”, “*shops close by*” (Wallgrün et al., 2014).

Three Arguments: In the case “near” is a comparative, the spatial relation takes more than one *LO* since the *LO* is evaluated with respect to a third object. Hence “nearer” means, an object is nearer to an object *Y*, than *Z* is and therefore is a relation taking three objects (see table 1).

There are different ways to compose these elements. Following Talmy (2000) a locative is treated as a relation between two spatial objects or entities. Alternatively, there is also a tradition within linguistics to treat “near” not as one relation, but to break the relation into two functions. The reason for doing so is to align the syntactic structure and its formal representation (for a discussion and overview see Zwarts (2017)). So instead of $near(LO, RO)$, the representation is $BE(LO, near(RO))$, whereas *near* is a function, which maps the *RO* to a relative location in space (place, location, region) ($RO \mapsto LOC$) and a second function, the general locating function *BE*, relating the *LO* with that relative location. The *BE* function is named after the copula “be” expressing the general relation between a figure and the relative location in space.

For a formal model of “near”, the NL expression is mapped to an abstract representation and usually the choice of the abstract representation depends on the linguistic properties chosen to

be formalized. Common abstract representations are relations ($near(LO, RO)$) (Grütter et al., 2010; Worboys, 1996, 2001), functions with two endpoints ($fs(LO, RO)$) (Bera & Claramunt, 2003), functions with two regions (Kracht, 2008), spatial triplets ($\langle near, LO, RO \rangle$) (Khan, Vasardani, & Winter, 2013), or predicates ($near(RO)$) (Bateman et al., 2010; Herskovits, 1986).

4.2 Properties of “Near”

Having defined “near” as a spatial nearness relation, we give a review of its properties. A cognitively adequate semantic model interpreting “near” needs to implement the following properties: *strictly positive* (sometimes also called minimality), *not necessarily transitive*, *not necessarily symmetric*, and *ignoring the triangle-inequality*. We describe the properties of the relation as a function with two endpoints, whereas O_n are spatial objects⁴.

Positiveness: Euclidean distance is positive, hence the range of a nearness function $fs(O_1, O_2) > 0$, if O_1 and O_2 are distinct, $fs(O_1, O_2) = 0$, if O_1 and O_2 are identical. In Worboys (1996) it is suggested that any distance function satisfies this condition.

This is a controversial point. In the case O_1 and O_2 are identical, then usually the preposition “near” is not applied, since the O_1 and O_2 are in contact with each other. However, in an empirical experiment by Logan and Sadler (1996), where subjects were asked in a production task to draw an X near a box, at least some of the participants drew their X touching the box. This might indicate that “at” (logically) entails “near” and we follow Levinson (2000) by arguing that the difference between “at” and “near” is the result of some added inference mechanism, which we will discuss in section 4.4. Nevertheless there are also approaches such as relative adjacency discussed in section 5.2.1.1, where the non-contact is part of the semantics.

Not-Necessarily Transitive: Transitivity is not necessarily given for a nearness function:

$(fs(O_1, O_2) \wedge fs(O_2, O_3)) = (fs(O_1, O_2) \vee fs(O_1, O_3))$. This is illustrated by Worboys (2001) with the following example: in the case of three locations: *Students’ Union, the Chapel, Academic Affairs*, where the Chapel is near Academic Affairs and the Student’s Union is near the Chapel, it does not necessarily follow that the Student’s Union is near Academic Affairs.

⁴Normally we call O_1 the LO and O_2 the RO

Not-Necessarily Symmetric: The distance between $fs(O_1, O_2)$ and $fs(O_2, O_1)$ is the same.

However, these two distances are not necessarily perceived as being of equal length, thus ignoring sometimes symmetry. Lee (1970) discovered that people underestimate the inward distance from the outskirts of the city to the city-center and overestimate the outward distance. However, Briggs (1973) and Golledge and Zannaras (1973) made the opposite observation as they noted that people underestimate the outward distance. Thus “near” is sometimes symmetric and sometimes not. Worboys (2001) calls this property of “near” weak symmetry.

Ignores Triangle-Inequality: The triangle inequality requires that the distance $fs(O_1, O_3)$ is less than or equal to the sum of distances between $fs(O_1, O_2)$ and $fs(O_2, O_3)$ i.e. $fs(O_1, O_3) \leq fs(O_1, O_2) + fs(O_2, O_3)$. It appears that distances are not necessarily perceived as being additive and it is possible that the sum of two distances $fs(O_1, O_2)$ and $fs(O_2, O_3)$ are perceived to be equal or smaller than the direct distance $fs(O_1, O_3)$ i.e. $fs(O_1, O_3) \geq fs(O_1, O_2) + fs(O_2, O_3)$.

Examples where the triangle inequality is ignored have been observed in cases of trip chaining (Brooks, Kaufmann, & Lichtenstein, 2008). The effect of trip chaining can be exemplified with the following example, where the electronic shop C is not considered to be near from home A. In the following example, the electronic shop C is not considered to be near from home A. However, if the trip is combined with a detour to a nearby grocery shop B, which, in turn, is considered to be near shop C, and needs to be visited anyway, the electronic shop C itself is then also seen as being near home A. Hence, in the case the visits of the shops B and C can be combined, the perceived distance to shop C decreases and the shop becomes near too. In this case shop C is only considered to be near from home A via the detour to shop B, but not via its direct way and thus ignore the triangle inequality rule in terms of the perceived distance assessment.

The properties of *not being necessarily symmetric, ignoring the triangle-inequality* (i.e., the “non-triangle-inequality”) ignore the Euclidean axioms (Bera & Claramunt, 2003; Qi, Shu, & Xu, 2006; Worboys, 1996). Thus, the conditions for metric spaces are not met (Huriot, Smith, & Thisse, 1989) and the relation “near” cannot be mapped directly into a spatial distance metric.

The *non-transitivity* of the nearness relation reflects the human inability to discriminate with

exact precision among things that do not differ much. Thus *tolerance* must be allowed for, which means that no precise distinction with respect to an observable property can be made (Minock & Mollevik, 2013). The effect of tolerance has been studied for the evaluation of options in consumer-choice preferences (Luce, 1956), the perception of approximate numbers (Dehane, 1997), ratio-dependent thresholds of discriminability (S. S. Stevens & Galanter, 1957), tolerance geometry (Roberts, 1973), and spatial indifference (Hubbard, 1978). To model tolerance, semiorders are used, where indiscriminability is not transitive⁵ (Luce, 1956). Since tolerant predicates (and relations) give rise to vagueness (van Rooij, 2011), the nearness relation itself is vague. This has the following consequences:

- The precision of a statement depends on the granularity of the observable property. The finer the granularity, the more differences matter and hence the smaller the *tolerance* and the less vagueness involved (for a discussion on granularity context and, vagueness see section 5.3).
- Some kind of indistinguishability will always be present due to incomplete information (see section 2.1) and “near” will always remain to some extent “vague”. This is because many spatial objects inherently are vague.

4.3 Semantics of “near”

As discussed in the previous section (section 4.2), there are two possibilities of formalizing “near”, either as a 2-place (“near(x,y)”) or 3-place relation (“near(x,y,z)”). In the case of “near” denoting a 2-place relation, the appropriate geometric representation is a region ($near_R$), and in case the denotation of “near” is a 3-place relation, then “near” is interpreted as a distance ($near_D$).

As noted earlier in section 4.1, in case “near” is used as a comparative, then the relation takes three arguments: *X is nearer Y than Z*. Treating the “nearness”-relation as a three-place relation “nearer(x,y,z)” has initially been proposed by van Benthem (1991) and subsequently, most notably, by Brennan and Martin (2002), and Nam (1995)⁶. Such a 3-place relation captures the comparative and relational aspect of “near” and formulates a nearness function with 3 argu-

⁵We assume that the findings of Ishikawa and Montello (2006) could also be formalized by semiorders.

⁶For an overview on “nearer(x,y,z)” see Zwarts (2017)

ments, the properties of which are as follows:

- Transitivity: $(fs(O_1, O_2, O_3) \wedge fs(O_1, O_3, O_4)) = fs(O_1, O_2, O_4)$
- Irreflexivity: $fs(O_1, O_2, O_1) = 0 \vee fs(O_2, O_1, O_1) = 0$
- Almost-connectedness: $fs(O_1, O_2, O_3) = fs(O_1, O_2, O_4) \vee fs(O_1, O_4, O_3)$
- Asymmetric on the 1st and 3rd argument: $fs(O_1, O_2, O_3) = 1 \wedge fs(O_3, O_2, O_1) = 0$ (Nam, 1995)
- Triangle inequality: $fs(O_1, O_3) \leq fs(O_1, O_2) + fs(O_2, O_3)$

To illustrate the triangle inequality for a 3-place relation, the relation can be depicted by two concentric circles around a common origin O_1 (see figure 1). Object O_2 is located on the inner circle, O_3 is located on the outer circle. Accordingly, the radius of the inner circle equals $fs(O_1, O_2)$ that of the outer circle equals $fs(O_1, O_3)$. It is easy to see that, independent of where O_2 and O_3 are located on the circles, the following holds: $fs(O_2, O_3) \geq fs(O_1, O_3) - fs(O_1, O_2)$. Hence we have: $fs(O_1, O_3) \leq fs(O_1, O_2) + fs(O_2, O_3)$ (Brennan & Martin, 2002; van Benthem, 1991).

Figure 1: Illustration of the Triangle Inequality for a 3-Place Nearness-Relation

One common strategy is to use “nearer” to model nearness, since it can be mapped into a metrical space. However, “nearer” cannot be reduced to “near” as pointed out by Brennan and Martin (2002, 2012), and Brennan, Martin, and Mihye (2004), since “nearer” cannot model “a is near b”. To overcome this problem they model nearness in terms of *impact areas*, which are based on regions. The main idea is to define nearness in terms of the intersection of the impact areas (IA) of two spatial objects. The axioms are as follows:

- (A1) For all objects o , $Near(o, o)$ (every object is near to itself)

- (A2) For all objects o_1 and o_2 , if $IA(o_1) \cap IA(o_2) = \emptyset \rightarrow \neg Near(o_1, o_2)$

Following these two axioms a family of “near”-relations are provided, whereas for us, the following “near” relations are relevant (definitions taken from Brennan et al., 2004):

$$s\text{-near-1}(o_1, o_2) = IA(o_1) \cap IA(o_2) \neq \emptyset$$

$$s\text{-near-2}(o_1, o_2) = o_1 \text{ belongs to } IA(o_2) \text{ or } o_2 \text{ belongs to } IA(o_1)$$

$$a\text{-near-3}(o_1, o_2) = o_1 \text{ belongs to } IA(o_2)$$

$$a\text{-near-4}(o_1, o_2) = IA(o_1) \text{ is a subset of } IA(o_2)$$

The implications are as follows:

- Positiveness: follows from axiom A1
- Not-necessarily transitive: the relations *s-near-1*, *s-near-2*, and *a-near-3* are not transitive. The relation *a-near-4* is transitive, but a limited transitivity (i.e. *downward monotone*)
- Not-necessarily symmetric: the relations *s-near-1* and *s-near-2* are symmetric, while the relations *a-near-3* and *a-near-4* are asymmetric.
- Triangle inequality: does not apply

Comparing the 3-place relation with the 2-place relation we make the following observations:

- The 2-place relation (i.e. a region, in this case IA), can adequately model the nearness relation. However, it lacks the relational properties to describe the distance of the different objects to each other (see our discussion on “near” with a third object in section 4.1).
- The 3-place relation (i.e. relative distance) cannot adequately model the nearness relation. Its relational properties, however, make it possible to compare the different objects to each other with respect to distance.
- Both relations do not model the limited transitivity adequately.

We address these observations as follows: In section 4.3.1 we discuss the ambiguity of “near”. There we discuss how “near” can be used to identify objects and their relational properties. In this case $near_D$ is more appropriate. Sometimes “near” can be used for spatial descriptions, and in this case, no such relational property is needed, the reason why $near_R$ is more appropriate. The problem of limited transitivity is regulated via tolerance and thus depends on the context. An adequate level of tolerance is discussed for granularity in section 5.3. The triangle inequality is discussed in section 6.

4.3.1 The Ambiguity of “near”

Zwarts (2017) observes that the preposition “near” sometimes behaves like a locative preposition with only topological properties such as “at”, “in”, or “on”. Sometimes “near” behaves like a distance preposition such as “above”, “behind”, “over”. An example where this different behavior can be identified are the two following sentences:

- Hotels near the Royal Opera House
- Let’s meet at the subway entrance near Big Ben

In the first sentence “near” is used as a nearness description to locate a target or targets (i.e. *hotels*) relative to a landmark (i.e. *the Royal Opera House*). To determine this denotation only knowledge of the place is needed and “near” denotes a relation between the RO (“*the Royal Opera House*”) and an acceptable region for “near” or a search domain for “near”, with the RO at its center, where the LO (“*Hotels*”) can be found (Levinson, 1996; Logan & Sadler, 1996). In our example we identify the neighborhood of Covent Garden as such an appropriate region containing spatial objects of type *hotel* as an adequate nearness description (see figure 2).

In the second sentence “near” is not used as a nearness description, but as a means to refer to an object. To evaluate the meaning of this sentence, different possible spatial objects need to be compared with each other to identify the appropriate reference. In this case “near” is a relation, which permits one to evaluate the distance between the spatial objects and the RO. For our example, this enables us to correctly identify and refer to “*the subway entrance near Big Ben*”. Consequently, all the different subway entrances need to be compared to each other with respect to their distance to “*Big Ben*” (see figure 3).

We identify an ambiguity with “near”, as “near” can denote a region ($near_R$) or a distance ($near_D$) and we identify this ambiguity with respect to the involved task (i.e. nearness description vs referential identification). The distinction is characterized as follows (Bateman, Tenbrink, & Farrar, 2007, p.178):

- *scene description (nearness description)* (localisation): speaker establishes where the objects or events are located using “near” and the denotation of “near” as a region ($near_R$) represented in terms of a 2-place relation. It is the preferred denotation of “near” for search tasks, as it answers to a *where-question* (near where is this located?)

Figure 2: Hotels near the Royal Opera House

Figure 3: Subway Entrances near Big Ben

- *referential identification* (speaker establishes the location of the objects with respect to each other using “near”, such as visiting an Indian restaurant “*is Ghandi’s nearer than Taj Mahal?*”, “*which Indian restaurant is the nearest?*”). In this case, a relational property of “near” is needed (“nearer to”), where spatial objects need to be compared to each other with respect to distance (see also “near” as relational knowledge (Couclelis, Golledge, Gale, & Tobler, 1987)). In this case the denotation of “near” is a (projective) distance relation ($near_D$) and an answer to a *which-question* (which object is near?).

According to Dale and Reiter (1995, p.223), a referring expression can convey two different kinds of information: *navigation* and *discrimination*. The goal of a *navigational* element is to bring the referent into the hearer’s attention and is used for *descriptions*. The goal of the *discriminational* element is used to rule out possible distractor objects when *identifying the correct referent*. The ambiguity of “near” arises because it can be used as either element.

In the spatial description task “*Hotels near the Royal Opera House*”, “near” is used as a *navigational* or *attention-directing* element and “near” denotes a region (see figure 2), which also can be represented as a set:

{The Savoy, Covent Garden Hotel, The Travel Lodge ... }

When “near” is used as a spatial description, a specification of the search space of acceptability is sufficient.

In the referential identification task “*Subway Entrances near Big Ben*”, the context set contains elements other than the intended referent. The goal of the discriminative element is to distinguish the intended referent (target) from other objects in the hearer’s focus of attention⁷. In this case “near” is used as a *descriptor* to discriminate the unwanted distractors for the actual referent.

In our example (see figure 3), the context set is a set of subway entrances:

{Speakers Green, Bridge Street, Boudiccan Rebellion}

Because we have a definite description, its uniqueness presupposition dictates that the expected answer needs to be a singleton: “*the subway entrance*”. The descriptor used to discriminate the unwanted distractor objects is “near” and singles out the other spatial objects which are not as

⁷See also Kelleher and Costello (2009), where they use Dale and Reiter (1995) generating referring expressions algorithm to select for an appropriate preposition given a visual dialog.

near. To rule out the unwanted objects, the different objects need to be compared with each other. To do so, the descriptor “near” is used to order the different subways with respect to a distance.

The ordering relation is induced via the relation ($nearer(x, \text{Big Ben}, z)$), where each triple contains the relative distance from a subway entrance with respect to another subway entrance to *Big Ben*:

$$\{ \begin{aligned} &< \text{Speakers Green, Big Ben, Bridge Street} > \\ &< \text{Speakers Green, Big Ben, Boudiccan Rebellion} > \\ &< \text{Bridge Street, Big Ben, Boudiccan Rebellion} > \end{aligned} \}$$

In this case, “near” is evaluated with respect to some distance metric which results in an ordered set. To satisfy the uniqueness presupposition, triggered by referential identification task, the first element from the list is chosen, which results in

$$\begin{aligned} &< \text{Speakers Green, Bridge Street, Big Ben} > \\ \Rightarrow &\text{Speakers Green is nearer to Big Ben than Bridge Street} \end{aligned}$$

The distinction between descriptor and descriptions can also be seen as a variation of “topology matters, metric refines” (Egenhofer & Mark, 1995), which says that purely qualitative topological relations are coarse-grained but are stage settings, whereas the metric properties provide the fine-grained distinctions when needed.

Various strategies making this ambiguity explicit are suggested in the literature. Denofsky (1976) calculates different thresholds for continuous ($near_D$) and discrete ($near_R$) cases. Crangle and Suppes (1989, 1994, p.194f) try to identify the distinction via different semantic compositional rules. The spatial ontology GUM-Space (Bateman et al., 2010) makes the distinction by introducing two distinct classes for proximity (projective and non-projective). The spatial annotation schema Iso-Space (Pustejovsky, Moszkowicz, & Verhagen, 2012) allows for a sentence to be encoded as a *metric* distance relation ($MLINK$) with the qualitative value of “near” (i.e. [The hot dog stand_{LO}] [$near_{MLINK}$] [Macy’s_{RO}]). In case of $near_R$ the entire phrase is encoded as a place ([near Harvard]_{RO}).

4.3.2 The Frame of Reference for “near”

$Near_R$ is not dependent on directional information since it describes how the two spatial objects relate to each other in terms of a region. In this case, the information of direction is irrelevant for this topologically defined spatial *nearness*-region, and thus $near_R$ is a topological preposition and not projective. Typical examples expressing a search space are as follows: *everything within reach is “near”* (Brennan, 2009) or *“The key needs to be near the house”*.

On the other hand, there are nearness expressions (including “near”, i.e. $near_D$), where *direction* plays a role when evaluating them and thus are projective.

Projective prepositions describe a region projected from the RO in a particular direction and the direction is interpreted with respect to a Frame of Reference (FoR). Usually, a FoR based on three types is used (Levinson, 1996) to interpret projective prepositions: *intrinsic*, *extrinsic* and *relative* (also sometimes called *deictic*) (see table 2), whereas the *extrinsic* and *relative* are external FoRs.

The relation between the different FoRs and nearness expressions are as follows:

- Intrinsic FoR: some entities have clear front and back sides, or left and right sides, which can be relied upon when forming spatial expressions. Typical cases are prepositions (*in front, beside, behind, above*), where *nearness* is restricted through a direction. An example is *“the bakery is behind the museum”*. In this example the denotation is as follows: *the bakery is within the conjunction of two regions: the nearness region of the museum ($near_R$) and in the region that is behind the museum ($behind_R$)*, whereas $behind_R$ is the region originating from the intrinsic “behind” the museum (Zwarts & Gärdenfors, 2015).
- Extrinsic FoR: the location is fixed within an extrinsic scheme of orientation (as, for example, in the points of the compass or some geotopological features such as uphill/downhill). In this case an acceptable distance for “near” is determined. Typical examples for $near_D$ depend on an extrinsic FoR, involving a path-following semantics (see also section 5.4.3 and Hedblom, Kutz, and Neuhaus (2015) for a formal definition of path). A path-following semantics of “near” relies on *distance* and *direction* information with respect to a network⁸. Furthermore, in a path-following semantics, a preposition is inter-

⁸See also the sections 5.2.2.2 and 5.4.3 for a discussion on distance in a network and direction.

preted as denoting paths or orientation, resulting in the denotation of “near” as *PAST*. In the example “*Geneva is rather near if you take the direct route via Lyon.*” at least two paths lead to *Lyon*, including the direct route, which passes *past* the location *Lyon*.

- Relative FoR: In case a relative FoR is present, the spatial property is evaluated with respect to a third object and not with respect to the properties of the *RO*. In case “near” is used to denote entities ($near_D$), direction plays a role, since “near” is evaluated with respect to a distractor object (see section 4.3.1).

Costello and Kelleher (2006) describes a setting for “near” with three spatial objects. These three spatial objects are the *LO*, the *RO* and a distractor *D*. The nearer *D* is to the *LO*, the less likely the *LO* and the *RO* are near to each other. Hence, the position of the distractor object *D* plays a role for evaluating whether the *LO* and the *RO* are near each other. In this case, the position of the distractor object sets the relative frame for $near_D$. This relative positioning is also relevant for setting the stage which is used to determine the granularity scale. This will be discussed in section 5.3.

Furthermore, experiments with discriminative tasks, such as ruling out distractors, have shown that orientation and direction is used as a cue when locating an object. This is in contrast to spatial descriptions, where no directional information is used (Ashley & Carlson, 2007).

Frame of Reference	Description	Example
Intrinsic	Proximity is determined by some inherent characteristics of the reference object.	“the bakery is behind the museum” (restricted $near_R$)
Extrinsic	Distance is determined by some external factors: arrangement of objects, travel time, or the costs involved.	“Geneva is rather near if you take the direct route.” ($near_D$)
Relative	Distance is determined by an external point of view or used figuratively.	“the one near me” ($near_D$): the speaker is the reference point (B. Tversky, Taylor, & Mainwaring, 1997). “tree near the house” ($near_D$) “the subway near Big Ben” ($near_D$)

Table 2: Frame of Reference (FoR)

Consequently, this means that “near” used for spatial proximity descriptions ($near_R$) is non-projective, whilst “near” used for object identification is projective ($near_D$). For instance, the sentence “The tree near the house” can be used to identify a particular tree in front of the house (projectively near), while other trees (non-projectively) are nearer and are part of the spatial description “trees near the house”.

4.4 Functional Relations for “near”

So far, we have defined the meaning of “near” in terms of geometrical relations (region, distance). However, research conducted within spatial language research shows empirical evidence that functional relation is also relevant for evaluating nearness (see section 2.2). As has been observed, if a functional interaction is present, it is assumed that the *LO* and *RO* are both functionally and spatially “near” (Coventry & Garrod, 2004). The spatial relationship of “near” is characterized as “possibly spatially interacts with the *RO*” In this section we discuss the motivation for introducing the notion of functional relation and discuss how the functional relation is applied to the two different interpretations of “near”: $near_R$ and $near_D$.

The question arises of the role of functional relation for larger spaces, as the literature on functional relations mostly investigates smaller spaces and less so larger scale space. Nevertheless, in the literature, the insights gained on functional relations are also transferred to larger space, for instance, as done by Bateman (2010a). Additionally, some studies within larger space implement a notion of functional relation, which we discuss in section 6.

For the spatial preposition “at”, the functional relation plays an important role. Usually “at” is thought to denote the touching between two spatial objects. However, the actual touching is only implied in the case there is a functional interaction between these two objects, such as in the case of “The man being at the piano”. When no functional interaction is present, such as in “the woman at the supermarket”, the “at” does not imply any touching, as “the woman” could also be standing near the supermarket. Hence, from a purely geometrical point of view, the touching between “the man” and “the piano” is not necessary. Only if the functional relation is given, and the two spatial objects interact with each other, is a touching assumed. In the case no interaction between the two spatial objects occurs, the possibility of the *LO* to be just near but not touching the *RO* cannot be excluded.

For “near”, a less strong notion of a functional-relation is assumed, meaning “possible to functionally interact with”⁹. This finding has been corroborated by Coventry and Garrod (2004), showing participants preferring proximity terms such as “near” over the preposition “at” in cases where no actual functional relation is present and the possibility of functional interaction

⁹For a different functional definition of “near” see Kemmerer (1999) and Bateman et al. (2010), where “near” and “far” is defined in terms of a functional contrastive opposite.

is not ruled out.

That is, “near” is not only evaluated with respect to its spatial representation but also with respect to its intended use. Similarly, Morrow and Clark (1988) characterize the verb “approach” (*become near*) as entering the region of interaction.

One way to model functional relations is by assuming an additional pragmatic layer in the modeling of spatial relations. One such additional layer, known as the inference mechanism for adding information, has been investigated as a means to add information to geometric meaning, which is known as Gricean inferencing (Grice, 1975). In the case of “near”, we have an entailment scale containing $\langle \text{at}, \text{near} \rangle$, where “at” (logically) entails “near” (Levinson, 2000). In this case, the proposition “the train is at the station” entails “the train is near the station”. However, due to Gricean inference rules, we know that when a speaker asserts a weaker item on a scale, it implicates that the stronger item is false. If the stronger item were true, the speaker would have said so. Because the proposition “the train is near the station” is weaker than the proposition “the train is at the station”, it means that the proposition “the train is at the station” is false. Therefore the statement “the train is near the station” actually means “the train is near the station and not at the station”. Hence, by assuming this inference process we have two elements under consideration: *what is actually at* (“at”) and *what is possibly at* (“near”), which corresponds to the functional relation notion of *what is actually functionally interacting with* and *what is possibly functionally interacting with*.

As noted by Bateman (2010b) and Zwarts (2017), this kind of pragmatic inference is probably related to the functional relation analysis, where nearness is investigated with respect to its functional relationship between its spatial objects. Depending on the task or goal, different objects or regions possibly interact with the agent and thus become “near”. Consequently, the models for “near” based on pragmatic theories and the functional relation analysis, where the extra-geometric properties are investigated, are connected to each other (see section 2.2). Furthermore, the interpretation of “near” becomes dependent on the task or goal of an agent, which we will discuss in section 6.

The functional relation is implemented with respect to the different denotations involved with “near”. In case of $near_D$ the functional relation is with respect to identified objects. That is, in this case the functional relation identifies the objects having the *possibility* of functionally

interacting with the *RO*. These objects then are possible candidates to be considered “near”. An example is the experiment conducted by Ferenz (2000), where participants considered only the couches facing the TV to be (potentially) “near”. The strategy of selecting appropriate candidate objects with respect to the task at hand (see section 6), is also implemented by pragmatic theories (see table 3).

In case of $near_R$, the functional relation is implemented with respect of a region of interaction (Carlson & Covey, 2005; Carlson & Van Deman, 2004; G. A. Miller & Johnson-Laird, 1976; Morrow & Clark, 1988). The region of interaction determines what is considered to be “near”, whereas different the property of this region are considered to be determined by the properties of the *RO* (see section 5.4) and the task at hand (see section 6). The strategy of generating an appropriate region of interaction with respect to the task at hand is also implemented by pragmatic theories (see table 3).

Functional Relation	Approach	Paper
Generates region	Information Theory	Worboys (2003)
Generates region	Bayesian approach	Ma, Kalashnikov, and Mehrotra (2008)
Generates region	Neuro-fuzzy	Barouni and Moulin (2015)
Generates region and selects candidates	Impact Area	Brennan and Martin (2012)
Selects candidates	Question Answering	Cai, Wang, and MacEachren (2003)
Generates region	Commuter Area	Cai, Wang, MacEachren, and Fuhrmann (2005), Minock and Mollevik (2013)
Generates region	Administrative Regions	Grütter, Helming, Speich, and Bernstein (2011)
Selects candidates	Map-Layer Selection	Grütter, Scharrenbach, and Waldvogel (2010)
Generates region	Ontology	Hahn, Fogliaroni, Frank, and Navratil (2016)
Generates region and selects candidates	Planning	Bateman (2010a, 2010b), Bateman, Hois, Ross, and Tenbrink (2010)
		Cai (2007), Cai, Wang, MacEachren, and Fuhrmann (2005)

Table 3: Different functional relations associated with “near”

5 Contextual Features of a Model of “near”

In this section we provide a synthesis of our integrative review. As discussed in section 3, we follow Turney (1996b) by distinguishing between primary and contextual features. Primary features are useful for classification even in isolation, whereas contextual features are only useful in combination with other features. A potential model assigns one or more context factors (CFs) to the qualitative or quantitative scales, where a scale can model “near” as a distance or a region (primary feature). In this case, a context factor is considered as a transformation on that particular scale (contextual feature). A cognitively adequate distance measure for “near”, however, needs to adhere to the properties described above (section 4)

The transformation depends on: (i) an underlying topology (flat or hierarchical, see section 5.1), (ii) the level of granularity (section 5.3), (iii) whether “near” is a distance or a region (i.e. what scale is appropriate, section 5.2) and whether the transformation affects both spatial objects or only one of them.

5.1 Topological Properties of the Domain

There are different ways to represent distance in space. The simplest model is to calculate the most direct distances on a map (i.e. distances in terms of “as the crow flies”), which has been criticized for not being cognitively adequate (Golledge & Zannaras, 1973). Instead, various topological representations of space have been suggested as an adequate representation for “cognitive maps” (Lynch, 1960). A topological representation usually includes places, paths, and regions that are linked together by relations such as connectivity, order, and containment.

Depending on the chosen model and cognitive assumptions, the topological map can either be flat, hierarchical, or mixed. The type of topology restricts the possible distance measures. In a strongly hierarchical network only regions can be compared to each other. Hence only topological or mereotopological scales are meaningful, while in a flat or partial hierarchical model, distance metrics are meaningful.

Flat Topology: In a flat topology only spatial information preserving topological connectedness is encoded. A flat topology is used to encode a street network (Byrne, 1979; Carbon, 2007; Helming, Bernstein, Grütter, & Vock, 2011) or a visibility graph (Montello, 1997; Nasar, Valencia, Omar, Chueh, & Hwang, 1985). In the case of vistas, a graph of the visible locations is created, where each vertex represents a location and each edge represents the visible connection (vista) between the vertices.

Strict Hierarchical Topology: In a strictly hierarchical topology different regions are stored in different branches. The nearness relation cannot be inferred directly by traversing the graph, but only from higher order geographic knowledge.

Possible predefined categories can be of geographical or cognitive nature. Countries and states (S. S. Stevens & Galanter, 1957; Vasardani et al., 2013) or administrative units (Grütter et al., 2010) are used as geographical categories. Districts or areas with different pop-

ulation density can be used as cognitive categories (Lynch, 1960). The effect of context is usually modeled via granularity and depending on the context a different granularity or scale level is chosen (see section 5.3).

Partial Hierarchical Topology: A partially hierarchical topology mixes flat and hierarchical topologies. Instead of using a predefined hierarchy, clusters are formed based on some properties of the flat topology. These clusters then form a hierarchy. Clusters, for instance, can be formed based on spatial memory. Places are clustered with respect to how quickly the subjects recall a place by using an ordered-tree algorithm (Hirtle & Jonides, 1985; McNamara, 1986; McNamara et al., 1989). Another possibility is by forming clusters with respect to geographic dispersion using a k -d tree (Roller, Speriosu, Rallapalli, Wing, & Baldridge, 2012).

In the literature the motivation for preferring a hierarchical model over a flat model is usually justified for cognitive reasons. However, there are also cases where a hierarchical model is chosen due to computational efficiency. One such case is Heath and Goodwin (2011), where they use a grid on a set of points as a hashing function.

5.2 Primary Features

Following Turney (1996b) primary features are useful for classification when considered in isolation. We identify the different distance measurements as primary features for “near” (see section 3) and classify them with respect to their scale (S. S. Stevens & Galanter, 1957).

Approaches modeling “near” on a qualitative scale (qualitative spatial reasoning, QSR) (section 5.2.1) are nominal and ordinal approaches, using the *connectivity relationship* or composition of the *distance relation* via an ordering preserving relation. Approaches modeling “near” on a quantitative scale (section 5.2.2) are based on an *interval* or *ratio* scale.

5.2.1 Qualitative Approaches to “near”

Qualitative approaches to “near” use qualitative distance measures (nominal or ordinal scales) to model nearness. These approaches have the advantage of not imposing a nearness metric and can thus deal with the problematic formal properties of the nearness relation (see sec-

tion 4.2). Although being cognitively adequate, such approaches are usually confined only to model $near_R$, but are not expressive enough to model $near_D$ (see section 4.1).

5.2.1.1 Nominal Approaches Nominal approaches are inspired by the theory of conceptual image schemata in cognitive linguistics (Lakoff, 1987; Talmy, 2000). The spatial relations are considered to be special cases of schemata, such as *containment*, *contact*, *path*, *scale*, or *cycle* (Lakoff, 1987).

Qualitative approaches investigate how these spatial relations can be represented in a way independent from the Cartesian two-dimensional space, from which they usually are extracted. For this purpose topological connectivity relations between points, lines and areas are used. More complex spatial relations can be derived from composition tables by what is termed qualitative spatial reasoning (QSR). An example of QSR is the region connection calculus (RCC) of Randell, Cui, and Cohn (1992). Topological relations as such are agnostic with respect to symmetry and the triangle inequality, which makes them suitable candidates for modeling the $near_R$ relation (see section 4.2).

Topological: Early experiments were promising, as they showed that these topological relations correspond well to NL spatial expressions (Mark & Egenhofer, 1994). This property is used to provide for a spatial model of NL (Bateman et al., 2010; Mark et al., 2007; Richter, Winter, Richter, & Stirling, 2013). The effect of context is mostly realized via granularity and its corresponding scale levels.

Alternatively in Bera and Claramunt (2003, p.355f) a relative distance measure, based on topological connectedness and named *relative adjacency*, is proposed. This distance measure takes as an input a dual graph derived from adjacency relationships and calculates the degree to which two regions are mutually distant from each other.

$$R(x_p, x_r) = d * Adj(x_p, x_r) + (1 - d) * \sum_{x_i \in Adj(x_p)} \frac{R(x_i, x_r)}{\#Adj(x_p)} \text{ where } 0 \leq d \leq 1$$

That is, the *relative adjacency* R of two regions x_p and x_r has the following elements: i) a binary adjacency relation Adj , an Adj set with the regions adjacent to x_p , and ii) a damping factor d balancing the importance of adjacencies and relative adjacencies.

Mereotopological: In Grütter et al. (2010) a mereotopological definition of “near” based on hierarchical levels is proposed. Two regions are “near”, if at least one of the conditions is

true: (i) both regions belong to the same common region next up in the hierarchy or (ii) one of the regions is immediately adjacent to the upper-level region. The scale level of the RO determines “near”. This approach can either be implemented with the RCC Calculus or with the algebra of sets (Özçep, Grütter, & Möller, 2012a). Depending on the context, an appropriate hierarchy level is chosen (see also section 5.3).

5.2.1.2 Ordinal approaches In ordinal approaches the objects are ranked. A quantitative (i.e. metric) scale is mapped to a weaker but still order preserving scale. For instance, a ranking for “near” is either achieved via relative distance (Gahegan, 1995) or via a fuzzy membership relation (Altman, 1994). Ordinal approaches are most commonly used to model the comparative forms of “near” (section 4.1).

Relative Distance (ordinal): The LOs are ranked with respect to (the inverse of) their absolute distance to the RO, where the first object is the nearest object. These ranked objects are then mapped to the NL categories “nearest”, “very near”, “near”, “not very near”, “not near” (the comparative forms of “near”), where the mapping itself depends on the context (Coventry & Garrod, 2004; Frank, 1992, 1996; Gahegan, 1995).

Yao and Thill (2005) use ordered logit regression (OLR) to predict near on an ordinal scale, given the distance (*absolute distance*) and context factors.

Fuzzy Membership: A region is generated by ranking the objects with respect to their distances using a fuzzy membership relation (μ_{MAX} & μ_{MIN}) in conjunction with a topological relation (Altman, 1994; Guesgen, 2002; Guesgen & Albrecht, 2000):

$$\mu_{MAX}(A, B) = MAX(\mu_{abs}(A, B), \mu_{rel}(A, B))$$

$$\mu_{MIN}(A, B) = MIN(\mu_{abs}(A, B), \mu_{rel}(A, B))$$

Most recently, however, a neurofuzzy approach has been suggested. Instead of defining a topological relation, a fuzzy membership relation is learned based upon a training corpus with contextual information inputs (Barouni & Moulin, 2015).

Composition of Intervals: In linguistics it is suggested to interpret directional prepositions in terms of paths (Lakoff, 1987; Talmy, 2000). One way of providing a formal model is by defining an algebraically structured set of paths that is thematically mapped to a spatial or geometric structure. Suggested algebraic structures are sets of intervals (Frank, 1992;

Hernández, Clementini, & Di Felice, 1995) or sets of vectors (Zwarts & Winter, 2000) where the size of the intervals depends on the context.

5.2.2 Quantitative Approaches to “near”

Quantitative approaches to “near” use distance measures on an interval or ratio scale to model nearness. A model of “near” with quantitative information is problematic, as the properties of “near” contradict metric properties (section 4.2). Since more information than only topological information is needed to model “near” as a distance ($near_D$), a quantitative approach is preferred over a purely qualitative approach. Using metrical scales for “near” is a nontrivial endeavor, as the conditions for metric spaces are not met by “near” (section 4.2). To make them cognitively more adequate, usually less expressive metrical scales (quasimetrics, semimetrics) are used and a contextually determined threshold needs to be determined¹⁰.

5.2.2.1 Interval Approaches Interval approaches are based on an interval scale and express distances, although with arbitrary zero value, which makes division impossible. Interval approaches to “near” are: *relative distance*, *Non-Euclidean Geometry*, and *similarity*.

Relative Distance (interval): The relative distance is the distance (interval) between two spatial objects with respect to a third object.

For instance, Worboys (1996) uses the following relative distance metric:

$$reldis(A, B) = \frac{dist(A, B)}{\mu_A}$$

where μ_A is the average distance from A to the other objects (Euclidean distance). Thus $reldis$ depends on the context, since the information of the distances between the other relevant objects determines the value of $reldis$. Furthermore, since $reldis$ is computed individually for each object pair, the distance metric is not symmetric and does not respect the triangle inequality (see section 4.2). This relative distance has been used by Brennan and Martin (2006). Another definition of relative distance is by combining ordinal relative with absolute distance (Gahegan, 1995; Hernández, 1994).

Relative distance is also used in a probabilistic model based on “spatial templates” (Logan & Sadler, 1996). Space is there partitioned into regions of ‘good’, ‘acceptable’, and ‘bad’

¹⁰For a discussion on different kinds of thresholds see Denofsky (1976).

(ordinal scale). The discrete ordinal scale, in turn, is extended to a continuous function, modeling a gradation of applicability (interval). Such a probabilistic function is very similar to fuzzy logic or fuzzy sets. In psychological or psychophysical settings, this kind of probabilistic relative distance is used to assess distance judgments of test subjects (Carbon & Leder, 2005; Couclelis et al., 1987; Gahegan, 1995; Kelleher & Costello, 2009; Logan & Sadler, 1996; Sadalla, Burroughs, & Staplin, 1980).

Non-Euclidean Geometry: Experiments have shown that some sort of metric information is used by respondents which does not strictly obey the metric axioms (Ishikawa & Montello, 2006). Some researchers hypothesize that “near” can be modeled by only using a restricted metric for “near”; geometric axioms are necessary for such a model.

Crangle and Suppes (1994) and Suppes (1991) discusses which geometry is appropriate for different prepositions. To model “near” they use *absolute geometry*, which is a weakened Euclidean geometry with the axiom of parallels removed. Since only congruence is required (for instance qualitative equidistance or ordering of qualitative distance, as investigated by Suppes, Krantz, Luce, and Tversky (1989)), such a geometry allows for *strict-positiveness* and *non-symmetry* (including weak symmetry). Unfortunately, this approach does not solve the problem of the triangle inequality, which still holds true in absolute geometry.

Zwarts and Gärdenfors (2015) model the space for prepositions in terms of polar coordinates, which allows for an egocentric representation. Points in space are ordered with respect to distance and angles resulting in a “near” region, which depends on the properties of the *LO*.

Besides determining a threshold, it is also hypothesized that the context determines possible axioms (Crangle & Suppes, 1994; Suppes, 1991).

Similarity: The so-called Tobler’s first law of geography says (Tobler, 1970, p.234): “Everything is related to everything else, but near things are more related than distant things.” Based on this “law”, it is hypothesized that “near” can be interpreted as a similarity relation between objects in the Euclidean plane. Additionally, it has been shown that the accuracy of similarity measurements is improved by including spatial relations such as “near” and “very near” for the calculation of semantic distances between geospatial concepts in a

conceptual space (Schwering & Raubal, 2005).

5.2.2.2 Ratio Approaches Ratio approaches to “near” are best explained in terms of index measurements such as *hunger* as an example (Luce, Krantz, Suppes, and Tversky (1990, sec. 22.7.5) and (Hand, 1996, p.458)). Instead of using the qualitative measure *hunger* directly, a proxy variable with the measure $\log(\text{amount of food ingested})$ is used. Such a proxy variable has the properties of (i) being well understood (ii) and easily measurable. Therefore we assume that ratio approaches to “near” (magnitude estimation, ratio estimation, absolute distance, probability density function, cost functions) can be seen as index measurements for “near”.

Ratio approaches to “near” are mostly based on *absolute distance*, which is either measured as a single distance (centroid to centroid) or as a *distribution* of distances (H. J. Miller & Wentz, 2003). In the latter case, the distribution is either represented as an *expected distance* (i.e. density function) or *average distance*. Distance distributions also serve as a basis for statistical approaches.

Magnitude Estimation: Magnitude estimation is a measure to scale sensory variables like loudness and brightness (A. Stevens & Coupe, 1978) and has been applied to determine *subjective distance*. In the case of magnitude estimation

$$Y = kX^n$$

the *subjective distance* Y is determined by the *objective distance* X , a *constant* k , and the exponent n . Empirical research has been performed to estimate the exponent n and hence the spread of the curve for “near” might change, given the context (Ekman & Bratfisch, 1965; Hourihan & Jones, 1979; Lundberg & Ekman, 1973; MacEachren, 1980; Radvansky, Carlson-Radvansky, & Irwin, 1995; Strzalecki, 1978; Wagner, 2006).

Ratio Estimation: The log-transformation of the magnitude estimation results in a ratio $\log(\text{Sub}D) = k + n \log(\text{Obj}D)$, having a linear property, and henceforth *SubD* can be determined via linear regression (Wagner, 2006). Another variant of ratio estimation is using hedonic house price data which also has a linear property (Lu, Charlton, Harris, & Stewart, 2014).

Absolute Distance: The most basic distance function is the Minkowski r -metric:

$$\text{dist}(x, y) = \left[\sum_{i=1}^n |x_i - y_i|^r \right]$$

where r is the Minkowski constant. In the case $r = 1$ we obtain the Manhattan distance (city block distance) and if $r = 2$, the Euclidean distance. Different contexts induce different thresholds on the absolute distance (Denofsky, 1976).

The use of absolute distance as a distance measure for “near” varies. Some use only absolute distance as a distance measure for “near” (Carlson & Logan, 2001; Delboni, Borges, Laender, & Jr, 2007; Dolbear, Hart, & Goodwin, 2007; Gapp, 1995; Hart & Dolbear, 2013; Roller et al., 2012; Tezuka, Kambayashi, & Takakura, 2001; Wallgrün et al., 2014; Yao & Thill, 2005), while others use it in combination with another distance measure (usually relative distance) (Barouni & Moulin, 2015; Brennan & Martin, 2006, 2012; Gahegan, 1995; Kelleher & Costello, 2009). Absolute distance serves also as a basis for statistical approaches (see *Relative Distance (ordinal)* and *Probability Density Function (PDF)*).

Probability Density Function (PDF): A statistical approach is to associate the nearness relation with a probability of the observer classifying the relation as true (e.g. “for a distance X , there is a specific probability Y , an observer will classify an object Z at this distance as ‘near’” (Denofsky, 1976)). These non-experimental models formalize the intuition that the nearer an object is, the higher its probability to be “near”. Depending on the assumption made concerning the nature of “near”, a different model is used: a Gaussian model based on simple frequencies from a corpus (hand crafted corpus (Denofsky, 1976), responses from a survey (Hall & Jones, 2012), web corpus (Schockaert, De Cock, & Kerre, 2008; Wallgrün et al., 2014), and a web n-gram corpus (Derungs & Purves, 2016)). Also more complex models, such as generalized linear models (Kienast, Degenhardt, Weilenmann, Wäger, & Buchecker, 2012), Gaussian mixture models (Skoumas, Pfoser, & Kyrilidis, 2013, 2014), or simple uniform Gaussian models (Ma et al., 2008), are used. Most of these models are based on absolute distance, with the exception of Hall and Jones (2012), where they also incorporate *relative distance (ordinal)* and *network distance (road connectivity)*.

Alternative ratio measures are based on a *cost function*. Such distance functions measure the required costs for moving across locations. Those costs can be measured in energy, money or time (H. J. Miller & Wentz, 2003, p.582). For “near” a time based cost function has been studied extensively in terms of *travel time*, *driving time*, *walking time*). Cost functions have the following

properties:

- Such distance functions are *symmetrical* and do not respect the *triangle inequality*, which classifies them as quasi-metrics. By introducing additional factors, such as horizontal and vertical factors or other contextual factors, will make these distance functions *non-symmetric*.

These properties make them more akin to the nearness relation (see section 4.2).

- In general, these distance measures do not measure a direct distance between two spatial objects. Instead, distance is measured by following a path. Usually a least-cost path relationship between these two objects is assumed and the cost function measures the efforts of movement.

Due to their properties, these measurements are considered to improve the modeling of “near” and hence are better predictors than Euclidean distance (Lu et al., 2014). Even though these distance functions are generally preferred, it should also be noted that “empirical work indicates that there is no ‘one-fit-all’ distance metric” (Lu et al., 2014, p.661). Distance measures based on a cost function are as follows:

Least-Cost Path: A simple cost-function has been proposed (H. J. Miller & Wentz, 2003): the distance relation is formulated in terms of a least-cost path through a cost field with terrain or land cover as its variables (Huriot et al., 1989). Thus, “near” can be defined as an object reachable below a certain cost.

Network Distance (flat): In general a network distance can be computed as soon as a topology can be constructed. The distance then is computed as the shortest path between the locations with respect to the amounts of vertices passed. A network is formed on a flat topology (see section 5.1) such as route networks (Allen, 1997; Brooks et al., 2008; Carbon, 2007; Gryl et al., 2002; Hall & Jones, 2012; Heath & Goodwin, 2011; Montello, 1997), transportation networks (Dao & Thill, 2009; Flushing, Gambardella, & Di Caro, 2013), or visibility graphs (vistas) (Montello, 1997; Nasar et al., 1985). In the case of vistas, a graph of the visible locations is created, where each vertex represents a location and each edge represents the visible connection (vista) between the vertices.

Network Distance (hierarchical): The network distance can not only be used to compute flat topologies but also hierarchical ones. Usually, the shortest path between the hierarchi-

cal elements is computed (Barsalou, 1985; Byrne, 1979; Cai, 2007; Couclelis et al., 1987; Heath & Goodwin, 2011; Kuipers, 2000; Marquet & Miralles-Guasch, 2014; McNamara, 1986; McNamara et al., 1989; Vieu, 1991). Alternatively, a distance measure is based on the comparison of distinguished features (matching-distance similarity measure (MDSM) (Rodríguez & Egenhofer, 2004)).

Travel Time: Travel time is a special case of general network distance where the distances in the network are based on the average interaction “cost” between two objects on a time scale. This is a special case of general network distance. Possible *cost measures* are travel time, driving time (Brennan & Martin, 2012; Cai, 2007; Helming et al., 2011; MacEachren, 1980; Montello, 1997), walking time (Marquet & Miralles-Guasch, 2014; Montello, 1997) or time on public transport (Kienast et al., 2012). In this setting, “near” means reachable within a certain time frame, or is considered as a form of temporal autocorrelation (Galton, 2009). To illustrate travel time, consider the following example of the towns Davos, Klosters, and Arosa in the Swiss Alps depicted in figure 4. In this example a flat topology is constructed with respect to the road-network in Switzerland where towns are represented as vertices and the connecting roads as edges. Since it is not possible to cross a mountain (unless there is a tunnel), such a network reflects the topographic properties of the Swiss Alps. While both towns are within the range of 11km from Davos, due to the topography of the Swiss Alps, the driving time to Arosa is 90 min and to Klosters 15 min. Now if a threshold of 15 min driving time is assumed ¹¹ following the road network, only Klosters can be reached within 15 min of driving time. Hence, given the assumption that 15 min driving time is “near”, only Klosters Dorf is “near” Davos, but not Arosa.

5.3 Granularity

Granularity is a concept which characterizes the level of detail and where spatial objects can be organized (via partitioning) into a hierarchical structure with respect to different granularities. The granularity of a scene is usually defined via a scale factor. One common scale factor has been suggested by Montello (1993) where he distinguishes between *figural*, *vista*, *environmental* and *geographical* space, which we introduced in section 2.2. Such a scale factor is based on an

¹¹For a discussion about different thresholds, see section 5.4

Figure 4: Different distance measures, Euclidean distance vs. driving time

1) OpenStreetMap is open data, licensed under the Open Data Commons Open Database License (ODbL) by the OpenStreetMap Foundation (OSMF)
 2) Boundaries by SwissBOUNDARIES © & Hillshade by swissALTI3D © 2019 swisstopo (5704000000).
 Reproduced by permission of swisstopo (JA10D118)

ordinal scale of different spaces.

Granularity itself is closely related to tolerance, as discussed in section 2.1. As we defined *tolerance* above, “near” is insensitive to small changes and these small changes themselves depend on the context. For instance, an acceptable tolerance in a figural space is different from a tolerance level in the geographic space. Hence granularity is a way to specify uncertainty and vagueness in the spatial domain. In general, it is assumed that the context dependency of “near” is captured by specifying a scale parameter and thus depends on the scale level of the distance measure discussed in section 5.2. We identify four different strategies concerning how to specify such scale parameters which we explain below and summarize in table 4.

Paper	Montello, 1993	Worboys, 1996	Gapp, 1995	Grütter et al. (2010)	Freundschuh, 1998	Gahegan, 1995
Scale	Ordinal Scale	Interval Scale	Interval Scale	Ordinal Scale	Ordinal Scale	Interval Scale
Scale Level	Figural		If marked with “from my point of view”			
Scale Level	Vista Environmental			Community Level County Level	Walking Driving	
Scale Level	Geographic	RelDis for Cities	Standard/Unmarked	Country Level	Flying	CVW

Table 4: Relation between FoR and scale

One strategy is to determine the scale parameter via a quantitative property of the *RO*. Investigated properties are, for instance, *the size of the RO* (Gapp, 1995), or *the number of inhabitants* (Worboys, 1996). Doing this will induce either an interval or ratio scale.

A second strategy is to set the scale parameter via the qualitative properties of the *RO*, based on administrative units of a country, as suggested by Grütter et al. (2010), Özçep, Grütter, and Möller (2012b). The administrative units are typically organized into a set of partially ordered partitions forming a partition chain. The units of the same partition share the same type and the nearness measure uses the information of the next upper level in the hierarchical structure. If the *RO* is a community, then a possible *LO* is either within the same county (a-priori near) or within the adjoining county. This will introduce an ordinal scale.

A third strategy is to specify the scale parameter via the expected interaction (functional relation). Freundschuh (1998) found that the geographic scale has an impact on how people estimate distance in everyday life. In fact, people tend to use (travel) time when assessing intermediate distances, ranging between 331 and 2,080 kilometers, while they use space to assess short distances (< 10 kilometers) and for very long distances (8,000 kilometers). This will introduce an ordinal scale.

A fourth strategy is to specify the scale parameter with respect to the *LO*, the *RO* and a third object. This strategy is used by Gahegan (1995), where he first establishes a bounding box from the current scene. Such a bounding box delimitates where to look for current possible objects and is called the “current viewing window”. Once $near_D$ is established, the scale is then determined in the same way as described in the first strategy. By doing so, this allows for establishing a deictic FoR (see section 4.3.2) and hence a distance measure on an interval scale.

5.4 Context Factors

Context factors (CFs) encode the information from the context. In this section we discuss the different CFs. Generally, it is assumed that “near” depends on various CFs representing the context. Determining nearness can be conceived as specifying a threshold on a qualitative or quantitative distance measure with respect to various context factors. A CF can transform the different nearness measurements. We observe that some CFs only affect distances ($near_D$) but not regions ($near_R$), or, in case a CF affects both region and distance, the effect might constitute itself differently (see table 5).

We also observe that the different CFs can be classified with respect to the following features: Some CFs can be associated only with one spatial object (either *RO* or *LO*), or with both spatial objects.

CF	Associates with	Effect on $near_R$	Effect on $near_D$	Example
Configuration	LO	N/A	Modifies Distance	Parallel vs. non parallel streets
Density of Objects	LO	N/A	Decreases Distance	Density of pubs, restaurants
Emotional Involvement	LO	N/A	Increases Distance	Affinity
Reachability	LO	Increases Region	Increases Distance	Street network, public transportation
Anchor Effect	RO	N/A	Modifies Distance	Distance in steps vs. travel time
Object Type	RO	Determines Region	Modifies Distance	Ordinary object, landmarks, POI
Population Density	RO	N/A	Modifies Distance	Urban vs. rural
Saliency	RO	Forms Regions	Increases Distance	Peculiar vs. no-peculiar building
Size	RO	N/A	Increases Distance	Museum vs. postbox
Barriers	Both Objects	Transforms Region	Decreases Distance	Track-lines, Rivers, Borders
Channels	Both Objects	Transforms Region	Increase Distance	Track-lines, Rivers, Borders
Cultural	Both Objects	Unclear	Unclear	Chinese vs. English vs. Yindjibarndi
Distractor Object	Both Objects	Transforms Region	Decreases Distance	Spatial objects, Landmarks
Direction	Both Objects	N/A	Changes Distance	North, East, South, West
Familiarity /Knowledge	Both Objects	Determines Granularity	Changes Scale	Hometown vs. traveling
Functional Interaction	Both Objects	Forms Regions	Increases Distance	Commuter flows
Hierarchical Effects	Both Objects	Determines Granularity	Increases Distance	San Diego-Reno effect
Memory	Both Objects	Shrinks Region	Modifies Distance	Memory vs. perception

Table 5: List of the CFs and their associated effects

5.4.1 Associating CFs only with the LO

The first element of the nearness relation $near(LO, RO)$ is the located object (LO). LO s are the target object of the nearness relation. Usually any kind of spatial object can serve as the LO , but it has been observed that in general the LO is smaller than the RO .

For instance, the sentence “the museum near the postbox” is considered to be odd. In this section we describe only the CFs, which exclusively associate with the LO and summarize them in table 7 and 8.

Configuration: The configuration of the streets around the LO has an influence on determining $near_D$. It seems that the way the streets align with the LO has an influence on the distance perceptions. In the case that the configuration of the streets is parallel, longer distances are still considered to be “near” compared to configurations with non-parallel streets (Byrne, 1979; Lloyd & Heivlyt, 1987; McNamara, Rump, & Werner, 2003; Montello, 1991). In smaller spaces the configuration is reflected with the direction of the objects. For instance in case of a couch facing the TV (Ferenz, 2000; Hund & Plumert, 2007)

Density of Objects: The availability of other objects of the same type has an influence on the distance of $near_D$ (Denofsky, 1976; Dolbear et al., 2007; Heath & Goodwin, 2011; Marquet & Miralles-Guasch, 2014). In the case of “Pubs nearby” or “Find me pubs near here”, the density of available pubs influences the distance. The more pubs available, the smaller the acceptable distance for $near_D$. Conversely, if there are fewer pubs available, the acceptable distance increases for $near_D$.

Emotional Involvement: Several studies have shown that subjects being emotionally involved with the LO has an effect on distance perception. Places with an affinity or special attraction appear nearer. Accordingly, the acceptable distance for “near” increases. Identified sources for affinity are political affiliation (Carbon & Leder, 2005; Ekman & Bratfisch, 1965; Lundberg & Ekman, 1973; Strzalecki, 1978), social affiliation (Carbon, 2007), popularity (Tezuka, Takakura, Lee, & Kambayashi, 2001), or attractiveness in terms of utility (Yao & Thill, 2005). In the case of political affiliation, Carbon and Leder (2005) identified a correlation between negative attitudes towards a political system and overestimation of the distance. Distance estimations are usually expressed in terms of Magnitude estima-

tion, or Euclidean distance. For this reason, these studies only take $near_D$ into account and not $near_R$.

Reachability: The reachability of the LO influences “near”. Especially for $near_D$, it has been observed that when related to a higher reachability, distance is perceived as shorter and thus increases the acceptable distance in general (Dao & Thill, 2009; Gahegan, 1995; Hahn et al., 2016; Kruijff, Lison, Benjamin, Jacobsson, & Hawes, 2007; Yao & Thill, 2005). For nearby recreation areas it has been observed that the higher the reachability, the more likely the visit (Kienast et al., 2012).

To model this effect a flat topology is used and network distances (Heath & Goodwin, 2011) or travel times are scaled with respect to reachability. Measures used for reachability are the intersection density (Downs & Stea, 1973; Lee, 1970; Sadalla et al., 1980) or the connected intersection ratio (Allen, 1997). Reachability can also be factored into Euclidean distance by using smaller thresholds for less reachable places (Brennan & Martin, 2006; Gahegan, 1995; Hahn et al., 2016; Schockaert et al., 2008; Yao & Thill, 2005).

In the case of $near_R$, reachability is measured with respect to connectedness in a hierarchical graph. In the literature many different ways to model the underlying topology, graph, or network are proposed (Bera & Claramunt, 2003; Brennan & Martin, 2006, 2012; Kuipers, 2000; Marquet & Miralles-Guasch, 2014; Schockaert et al., 2008).

5.4.2 Associating CFs only with the RO

The second element of the nearness relation $near(LO, RO)$ is the reference object (RO). Its role is to specify the location of the LO . The RO serves as an anchor point for the nearness relation and determines the region of interaction. ROs are of type object (buildings, landmarks elements of the surrounding, streets) or they are geographical environments (areas, field names, environmental frameworks (Mainwaring, Tversky, Ohgishi, & Schiano, 2003)). In this section we describe the CFs associated with the RO (summarized in table 7 and table 8).

CFs that only associate with the RO usually factor in some properties of the object (e.g., *object type*, *salience*, *size*) and exploit the *anchor effect*:

Anchor Effect: The anchor effect is a cognitive bias and is observed when a person relies excessively on a single piece of information (A. Tversky & Kahneman, 1974). A study by

LeBoeuf and Shafir (2009) shows that participants systematically make wrong distance estimates when the scale unit is already provided. This bias can be explained by the anchor effect: Participants anchor on the “here” or “now” and make adjustments to accommodate the provided unit of distance. However, Burigo and Coventry (2010) were unable to reproduce the anchor effect for “near”.

So far, the anchor effect has only been observed for $near_D$ but not for $near_R$.

Object Type: It has been noted that the type of object also has an effect on the nearness relation (Denofsky, 1976; Ferenz, 2000; Grütter et al., 2011; Hernández, 1994; Schockaert et al., 2008; Xu & Klippel, 2012) (see table 6).

Object Type	Example
Movable Object	near the bicycle, near the postbox
Building	near the post-office, near the radio tower
Area	near the parking area, near the market
Field Name	near Shoulder of Mutton, near Six Acres
Stationary Objects	near the letter-box, near the public phone booth
Landmarks	near Big Ben, near Westminster
Environmental Framework	near the surrounding buildings, near the watershed “Willamette River Basin”
Streets	nearby Oxford street, near Store street

Table 6: Different Types of Objects

In general, the more prominent an object is, the smaller the distance judgment, which results in higher $near_D$ thresholds. In the case of very prominent objects, such as landmark objects or culturally important objects (Tezuka, Kambayashi, & Takakura, 2001), this effect is thought to be very pronounced.

Additionally, experiments have shown that landmarks introduce an asymmetry on the nearness relation. People judge the distance from an ordinary structure to a landmark to be shorter than vice versa (Couclelis et al., 1987; McNamara & Diwadkar, 1997; Sadalla et al., 1980).

Population Density: How densely an area is populated has an influence. “Near” is interpreted differently within an urban context than in a rural one. For $near_D$ it has been studied how the population density affects expected travel time (Hubbard, 1978; MacEachren, 1980; Marquet & Miralles-Guasch, 2014). An illustrative example is provided by Brennan and Martin (2012) where it has been noted that “near” in Australia can mean up to a travel day. Additionally, Worboys (1996) defines a relative distance metric between larger

cities by factoring in the population size of each city. However, comparing US cities, an experiment using a web co-occurrences corpus has shown that the variance of the median of Euclidean distance between toponyms varies significantly: for New York City 4km, Los Angeles 19km, and for San Francisco 25km (Derungs & Purves, 2016).

Salience: A well known principle from Gestalt psychology is the grouping of proximal objects which are then perceived as such. Objects of similar shape, size or color are more likely to form groups. For instance, smaller objects get a higher salience rating than bigger potential candidates, having the same feature of visual salience. Consequently, the size of the different possible objects needs to be compared to determine visual salience. This can be achieved by comparing the horizontal and vertical extensions (Gapp, 1995) or the convex hull (Barclay & Galton, 2013; Gapp, 1995).

Size: The size has an influence on both $near_D$ and $near_R$. The bigger the object, the smaller the distance judgment (Barouni & Moulin, 2015; Burigo & Coventry, 2010; Carlson & Covey, 2005; Couclelis et al., 1987; Denofsky, 1976; Ferenz, 2000; Gapp, 1995; Hernández, 1994; Kelleher & Costello, 2009; Liu, Guo, Wieczorek, & Goodchild, 2009; Wallgrün et al., 2014; Yao & Thill, 2005). In the case of $near_R$, since a bigger RO induces a larger search space (Gapp, 1995), the size has an influence on determining an acceptable nearness region.

5.4.3 Associating CFs with both Objects (LO & RO)

A wide class of CFs have an influence on both spatial objects (LO and RO). These CFs can either have an effect on a distance threshold ($near_D$) or determine a region ($near_R$) (summarized in table 9 and 10).

Barriers: *Barriers* are elements that inhibit a direct movement and have a direct influence on “near”-judgments (Baird, Wagner, & Nomu, 1982; Delboni et al., 2007; Hirtle & Jonides, 1985; Lynch, 1960; Montello, 1997; Wagner, 2006). Typical barriers are either natural or artificial objects, such as track-lines, rivers, mountains, or borders. For instance, in our example of the towns Davos, Klosters, and Arosa (figure 4) discussed in section 5.2.2.2, where the mountains in the Swiss Alps act as barriers. In this case a barrier prevents Arosa being near Davos.

In the case of nearby recreation areas a barrier effect has been observed for motorways,

ridges, and valleys (Kienast et al., 2012). Barriers can also serve as distractor objects and therefore the *LO* and *RO* are evaluated with respect to a barrier (see van der Zee et al. (2009) for an empirical study within figural space).

However, the effects of barriers are not universally accepted. For instance, O'Keefe (1996) doubts the effect of barriers. He argues that the following sentence is perfectly sound: "The house is nearby, but it will take a long time to get there since we have to go the long way around."

In the case of $near_D$, the barrier influences the distance perception. That is, estimated distances are greater when a barrier between the objects is present. As a consequence, objects that are previously considered as near to each other will not be perceived as such anymore. Hence, if a threshold has been determined for "near" (Denofsky, 1976), then this threshold is lower in the presence of a barrier. This has been observed for almost all distance measures (see table 9 and 10). Additionally, it has been observed that this effect is additive. Thus, the more barriers present, the higher the distance estimations (Thorndyke, 1981).

In the case of $near_R$, the effect of a barrier in a topological model is the interruption of the connection between the two objects (Hernández et al., 1995). Consequently, the properties of the *RO* are transformed with the result that the *LO* is not contained in the region anymore.

Channels: Channels are objects along which persons move (Lynch, 1960). An object which serves as a barrier can also serve as a channel at the same time. This is because barriers cannot be crossed, at least not easily, but one can still move along them. The area along a barrier is then considered to be "near". Examples for channels are walkways, transit lines, canals, railroads (Gryl et al., 2002), and borders.

In the case of $near_R$, a topological transformation is assumed and a network along the barrier is formed (Brennan & Martin, 2012; Gryl et al., 2002). For $near_D$, an area (influence area, impact area) along the barrier needs to be defined by a threshold, and within this area the condition for nearness must be satisfied (Denofsky, 1976).

Cultural: Mark et al. (2007) argue that "near" is culture dependent by looking at Yindjibarndi, an Australian Indigenous language. However, when comparing "near" used by Chinese

or English speakers (Xiao & Liu, 2007) or by German and English speakers (Hall & Jones, 2012), this finding could not be replicated. These two studies indicate the dominance of local knowledge over cultural effects. Furthermore, Yao and Thill (2005) found no significant factor in qualitative proximity modeling among different ethnic groups within the United States. The only exception is African Americans, accepting trip destinations to be further away compared to the other ethnic groups.

Distractor Object: A distractor object is an object intervening between two spatial objects (Ashley & Carlson, 2007; Costello & Kelleher, 2006; Denofsky, 1976; Hund & Plumert, 2007; Kelleher, Kruijff, & Costello, 2006). If two objects are near each other and a third object is introduced nearer to the *RO* than to the *LO*, then the two initial objects are no longer near. The same effect occurs by introducing a more prominent object than the *RO* (for instance a landmark object) (Tezuka, Takakura, et al., 2001). Experiments have shown that the relative distance of the distractor object to an additional object has no effect on “near” (Burigo & Coventry, 2010).

This effect has only been observed for $near_D$, but not for $near_R$. We assume that the effect of the distractor object is coupled with distances and directions (see also section 4.3.2, where we discuss, how the distractor object introduces directionality)

Direction: On the one hand, there are studies arguing that direction — cardinal direction (North, East, South, West), relative direction (left of, right of, in front of, etc.) — have no influence on the interpretation of “near” and thus consider “near” to be non-projective (Carlson & Kenny, 2006; Coventry & Garrod, 2004; Logan & Sadler, 1996; Mainwaring et al., 2003; Radvansky et al., 1995). Furthermore, there is also a wide variety of formalisms which try to incorporate direction into a representation of “near” (Frank, 1992, 1996; Kelleher & Kruijff, 2005; Kelleher et al., 2006; Skoumas et al., 2013; Zwarts & Gärdenfors, 2015; Zwarts & Winter, 2000). Sometimes it is argued that other factors, such as reachability (Gahegan, 1995; Yao & Thill, 2005), presuppose directionality.

On the other hand, there are experimental studies arguing that the direction has an influence on the nearness relation and thus consider “near” to be projective (Carbon, 2007; Carbon & Leder, 2005; Carlson & Covey, 2005; Carlson & Van Deman, 2004; Couclelis et al., 1987; Hourihan & Jones, 1979; Kuipers, 1978, 2000; Lee, 1970; Lundberg & Ekman,

1973; A. Stevens & Coupe, 1978) or experimental studies within table-top space (Carlson & Kenny, 2006).

Additionally experimental studies showed that functionally related objects are considered more likely to be near when they face each other (Coventry & Garrod, 2004; Ferenz, 2000; Hund & Plumert, 2007).

We attribute this controversy about the ambiguity of $near_R$ and $near_D$ to the fact that direction only plays a role when “near” is used for object identification but not for scene description (see our discussion in the sections 4.3.1 and 4.3.2). To our knowledge there are very few studies controlling for this difference (i.e. Ashley and Carlson (2007)).

Familiarity/Knowledge: In the case of $near_D$, the effect of familiarity is a widely researched topic. In general it is assumed that the more familiar the area under consideration (i.e. the surrounding area of the RO and the route to the LO), the smaller the distance assessment, although the factual distance is higher (Barouni & Moulin, 2015; Yao & Thill, 2005). Also Gapp (1995) reports that an already mentioned object has an effect on the perceived distance for “near”.

The effect of knowledge on “near”, however, is disputed. In an early study by P. F. Fisher and Orf (1991) a correlation between “near” and knowledge could not be identified. Additionally, in most experimental studies, evidence is limited because participants are only tested in environments familiar to them (Gahegan, 1995; Worboys, 2001; Yao & Thill, 2005). Similarly, Brennan and Martin (2012) only consider domain experts familiar with the area for the task of determining an impact area. The only study that compares nearness judgments of familiar and less familiar environments is conducted by Hall and Jones (2012). In this study they discovered that the knowledge of an area is important for determining the scale factor and qualifying $near_D$ as *road-distance*, *banded distance*, or *angle distance* (see also section 5.3). This observation aligns well with the prototype effect also observed for “near”. With limited information available the subject relies more on categorical information (see “Hierarchical Effects” below). On the other hand, with more information available, the subject relies less on categorical information.

Hierarchical Effects: Hierarchical effects are observed for “near”. A. Stevens and Coupe (1978) observed the so-called San Diego–Reno effect. From memory, subjects in San Diego, Cali-

ifornia were asked to indicate the direction to Reno, Nevada. Most subjects indicated that Reno is to the northeast. This disagrees with the fact that Reno is in the northwest of San Diego. Since Nevada is east of California, the subjects must have drawn the inference that every city in Nevada is east of every city in California. Since this effect could be replicated for different cities in the US, Stevens and Coupe conclude that it is not the cities that have been put into spatial relation but rather the regions containing the cities (i.e. the upper element in the hierarchy). Hence, spatial relations in hierarchies, including “near”, are not evaluated via distance but via hierarchical information. Accordingly, Maki (1981) observed that in the case that two places belong to two different regions, the distance between the two places is systematically overestimated.

The implications are as follows: for $near_D$ a partial hierarchical topology is assumed, since points need to be evaluated with respect to their hierarchical properties. If both spatial objects are confined within the same region (i.e., belong to the same hierarchical element), the distance is perceived as shorter (Hirtle & Jonides, 1985), which results in increased distances for $near_D$. Conversely, if the two objects do not belong to the same region, then the distances are perceived as longer, resulting in decreased distances for $near_D$ (McNamara, 1986; McNamara et al., 1989).

For $near_R$ a strict hierarchical topology is assumed (see section 5.1), where only the hierarchical elements can be compared to each other. For instance, it is checked if both spatial objects are within the same region (i.e., belong to the same upper level category). The role of context in this case is to determine the granularity (i.e., scale level in the hierarchy, see also section 5.3) (Grütter et al., 2010; Maki, 1981; Worboys, 2001).

Memory: Whether “near” is evaluated from memory or from perception influences the nearness judgments. In general, it is assumed that the more a judgment relies on memory, the less reliable this judgment, as it relies more on categorical information (McNamara et al., 1989). In turn, judgment from memory is more schematized, which should reduce the cognitive load. For instance, ROs of the same type, being nearby, or having a functional interaction, are grouped together (B. Tversky, 2003). Additionally, it is assumed that memory has an influence on determining the frame of reference (McNamara et al., 1989; McNamara et al., 2003).

In the case of $near_D$, it is assumed that memory is a source for the asymmetry in distance estimation (McNamara & Diwadkar, 1997; Sadalla et al., 1980). Moreover, memory has an effect on distance estimation (Byrne, 1979; Montello, 1993; Taylor & Tversky, 1996; Thorndyke, 1981; Wagner, 2006), where memory can either increase or decrease distance estimation. One hypothesis is that memory acts as an amplifier for the other CFs.

For $near_R$, B. Tversky (2000) observes that memory has a shrinking effect on the region.

Weather, Age, Dynamic Elements In Minock and Mollevik (2013), this idea is extended by factoring in weather and time of the day. Furthermore, the age of the agent plays a role (Kienast et al., 2012).

The findings of this section are summarized together with the distance measures of section 5.2 in table 7, 8, 9, and 10.

6 Contextualized Tasks for “near”

From a cognitive or intelligent *agent* point of view, context can be distinguished between external and internal factors (Cai, 2007; Tamminen, Oulasvirta, Toiskallio, & Kankainen, 2004). External factors are factors dependent on the space (see section 5), whereas factors such as *tasks* and *goals* are agent internal. In this section, we discuss the interpretation of “near” depending on agent internal factors, such as the task.

In one example case, the interpretation of “near” depends on two different tasks (Cai, 2007): *buying non-organic food* and *buying organic food*. If the task is to *buy non-organic food*, the grocer’s shop within 10 minutes walking distance is “near”, whereas the organic market within 15 minutes driving distance is considered to be “not near”. If the task, however, is to *buy organic food*, then the organic market within 15 minutes driving distance is considered to be “near”.

In this section we provide an overview of the different “nearness”-tasks (summarized in table 7 and 8) and discuss how tasks induce different functional relations. Subsequently, we discuss how the different tasks can be implemented via Turney’s contextualisation framework (Turney, 1996a) by means of *contextual weighting* and *contextual classification*.

When an agent has a task at hand, very often that agent needs to evaluate the task with respect to “nearness”. As discussed in section 2.2, the pragmatic theories exploit the fact of how

“near” is embedded within a communication process.

Within such a setting, following Bateman et al. (2007), we identify the following generalized tasks: *scene description*, *referential identification*. As discussed in section 4.3.1, we identify a *scene description* with $near_R$ and a referential identification with $near_D$ (descriptor). Besides these two tasks (Bateman et al., 2007), we identify *route description* as a task involving near, where the combination of scene description and referential identification is possible. Furthermore, we also make the same observation, where scene description and referential identification can be combined together in other complex tasks, such as land-use planning:

- *route description* (speaker is instructing the hearer to reach a goal location: “*in the round-about near Octon Cross, turn right and follow the road until approaching (getting near) Eddsfield Airfield*”).
- *combination of the two tasks (land-use planning tasks)*: When planning a *waste dump location*, questions such as the following need to be addressed: “*Identify if a possible waste dump location is not too near to other relevant spatial entities.*” (Brennan & Martin, 2012)

We will use this characterisation to summarize the different nearness-tasks in table 11.

Looking at table 11, we can make the following observations: the *task at hand* and the relevant features are functionally related to each other. Depending on the task, different features (primary features and CFs) are relevant. In general, it is assumed that the external and internal context factors are related to each other with respect to relevancy, which implies that for each task different CFs are functionally relevant. The “expected interaction” assumes that, for the task at hand, only relevant CFs can be used.

Hence, depending on a task, different relevant contextual information becomes more or less appropriate. We here have a two step process for actually defining “near”, which is then evaluated with respect to the selected features (the two levels correspond to the semantic level and the functional level discussed in section 4).

Turney (1996a) calls this two step process contextual classification. A specialized classifier is selected from a set of classifiers, based on the CFs. Then the specific classifier is applied to the primary features.

In order to build a specialized classifier for “near”, we identify two strategies in the literature of pragmatic models. The first strategy is to select an appropriate region, considered to

be potentially “near”, the second strategy is selecting appropriate candidates (see table 3 in section 4.4).

The task also has an influence on the interpretation when “near” ($near_D$) is associated with multiple objects. Then it is possible to construct a group-based reference with multiple objects grouped together. Ingene and Ghosh (1990) compares the distance between the starting point and the first stop to the distance of the combined trip. They observed that the smaller the difference between the two distances, the more likely a consumer is to chain the trip. Brooks et al. (2008) discovered that a consumer not only tries to minimize the total travel distance but also tries to minimize subjective travel costs. For similar routes, objects are grouped by function: commercial buildings grouped together with other commercial buildings, i.e. university buildings grouped together with other university buildings (Hirtle & Jonides, 1985).

The effect of the task can also be quite substantial. For instance, as discussed earlier in section 5.4.1, the sentence “the museum near the postbox” is considered to be odd. However, in case of a wayfinding task, with the starting point at the postbox, and two museums in the precondition set, will successfully get you to “the museum near the postbox (as opposed to the other museum across the street)”.

This has a consequence for the different tasks. For a task at hand, the functional relation says which objects or regions are considered to be potentially “near”. Hence, the functional relation is definitional for “near”: it defines what is potentially near or considered to be near. For a specific task, however, the different context factors are relevant (i.e. “near” itself is functionally indeterminate) and “near” is evaluated with respect to these relevant context factors. Therefore, the CFs are information relevant to the success of the goal.

7 Conclusion and Outlook on Future Research

As this review shows, there is a large body of literature investigating different CFs for “near”. We attribute this large body of literature to the interdisciplinary nature of the problem as well as the manifold situations in which the expression “near” is used. Sources for the broad range of the identified CFs arise from the properties of the related or co-occurring spatial objects. Furthermore, we also attribute the availability of a wide range of publications to the many

ways in which “near” can be represented via different distance measures.

The literature usually considers a CF in isolation. Only the correlation and interaction effects of such a CF with a distance measure is investigated. However, if the task is to predict nearness, a nearness model needs to be formulated. In this case, only considering a CF in isolation is not adequate. Instead, a set of CFs needs to be considered. For example, in the case of “near for a daily commute”, the sizes and types of the objects in combination with the available mode of transportation and reachability have to be taken into account. In agreement with the literature, the CFs alone cannot predict “near” but depend on a distance measure.

Thus, to formulate an adequate model for “near”, a methodology is needed, allowing the comparison of the different CFs with respect to a distance measure in a systematic way. To do so, we combine the methodology of the integrative review with Turney’s contextualisation framework.

The result of our integrative review is presented as a structured table of features: 7, 8, 9, and 10. These tables allow us to compare the different features for “near” and lay the foundation for formulating a formal model for “near”.

For each table the first dimension represents the primary features and is identified with the qualitative and quantitative distance measures. These measurements have the property of being meaningful in isolation and can be used to predict nearness. The second dimension represents the contextual features and is identified with a set of CFs. CFs have a contributing effect on determining “near”, but have no meaning in isolation. For the identified CFs, we determine their transformative effect on a distance measure and their ability to determine different thresholds.

With respect to the semantics of “near”, we identify two different possible denotations: “near” can be used as a spatial description or as a spatial descriptor. Our review shows that “near”, as a spatial description can be represented sufficiently well by regions ($near_R$). In this case qualitative structures are adequate. “Near” as a spatial descriptor ($near_D$), on the other hand, is usually better represented by quantitative or semi-quantitative structures. Hence, we discovered that an adequate representation depends on the semantics of “near”. However, this point is controversial. Very often it is argued that spatial relations in general, and the nearness relation in particular, are represented cognitively more adequately as qualitative relations, where regions and lines are treated as primitives. This is in contrast to quantitative representa-

tions where points and distances between points are treated as primitives.

With respect to CFs and the $near_R/near_D$ ambiguity, we observe a dependency. CFs with qualitative properties (e.g., salience of a spatial object, familiarity with an area or a route) can be associated with both $near_R$ and $near_D$. In the former case, a region is transformed and in the latter case it is a distance that is transformed. However, CFs with quantitative properties (e.g., population density, density of objects), are only associated with $near_D$.

This has an implication for the adequate representation in terms of the semantics of spatial prepositions (Cohn & Renz, 2008; Zwarts & Gärdenfors, 2015). The scale level of the CFs, associated with the preposition, also need to be taken into account. Until now, the link between properties of the CFs and an adequate representation has not been explored in the literature.

Our study also reveals a functional relation between a task at hand (e.g., *work*, *shopping*, *school*, or *land use planning*) and appropriate features. In the literature, this relation is only implicitly assumed. Our contribution is to make this relation explicit by introducing a two-step process. First, the relevant features are selected for the given task and second “near” is evaluated with respect to the selected features.

To the best of our knowledge, however, no study exists where the relation between the kind of task performed and appropriate features are investigated empirically for larger spaces. Table 11 can be used as a starting point for an empirical study, determining the relevant features (distance measures and CFs) for a given task. Consequently, the results of this study can be used to build specialized nearness classifiers for specific tasks.

Acknowledgement

We would like to thank Maria Elena Vargas Amado for the drawing and design of the maps. This work is partly funded by the Swiss National Science Foundation (SNSF) Project No. 200021_130151/2 and the Swiss Federal Office for the Environment (FOEN). We would like to thank the editor and the four reviewers for their suggestions and critical comments that helped to improve this paper.

Table 7: CFs associated with one spatial object on a qualitative scale

Scale	Distance Measure	Associating CFs with the LO				Associating CFs with the RO				
		Configuration	Density of Objects	Emotional Involvement	Reachability	Anchor Effect	Object Type	Population Density	Salience	Size
Nominal	Topological				Bera and Claramunt (2003)					
	Mereotopological						Grütter, Helming, Speich, and Bernstein (2011)			
Ordinal	Relative Distance (ordinal)	Hund and Plumert (2007)		Yao and Thill (2005)	Yao and Thill (2005)		Gahegan (1995), Hall and Jones (2012), Yao and Thill (2005)			
	Fuzzy Membership				Schockaert, De Cock, and Kerre (2008)			Robinson (2003)		
	Composition of Intervals							Hernández, Clementini, and Di Felice (1995)		
Interval	Relative Distance (interval)			Carbon and Leder (2005)	Brennan and Martin (2006), Gahegan (1995)		Gahegan (1995), Kelleher and Costello (2009)	Worboys (1996)		Couclelis, Golledge, Gale, and Tobler (1987), Hernández (1994), Kelleher and Costello (2009)
	Non-Euclidean Geometry									Crangle and Suppes (1994), Suppes (1991), Zwarts and Gärdenfors (2015)
	Similarity									

Table 8: CFs associated with one spatial object on a quantitative scale

Scale	Distance Measure	Associating CFs with the LO				Associating CFs with the RO				
		Configuration	Density of Objects	Emotional Involvement	Reachability	Anchor Effect	Object Type	Population Density	Salience	Size
Ratio	Magnitude Estimation			Ekman and Bratfisch (1965), Lundberg and Ekman (1973), Strzalecki (1978), Wagner (2006)	Wagner (2006)			Houriha and Jones (1979), MacEachren (1980), Wagner (2006)		McNamara and Diwadkar (1997)
	Ratio Estimation					Burigo and Coventry (2010)		Liu, Guo, Wiczorek, and Goodchild (2009)		Burigo and Coventry (2010), Carlson and Covey (2005), Liu, Guo, Wiczorek, and Goodchild (2009)
	Absolute Distance	Ferenz (2000)	Denofsky (1976), Dolbear, Hart, and Goodwin (2007), Pareti and Klein (2011)	Houriha and Jones (1979), Yao and Thill (2005)	Brennan and Martin (2006, 2012), Gahegan (1995), Hahn, Fogliaroni, Frank, and Navratil (2016), Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012), Schockaert, De Cock, and Kerre (2008), Yao and Thill (2005)		Brennan and Martin (2012), Canneyt, Crommenlaan, Schockaert, Laere, and Dhoedt (2012), Delboni, Borges, Laender, and Jr (2007), Denofsky (1976), Ferenz (2000), Gahegan (1995), Gapp (1995), Hall and Jones (2012), Schockaert, De Cock, and Kerre (2008), Tezuka, Kambayashi, and Takakura (2001), Wallgrün, Klippel, and Baldwin (2014), Yao and Thill (2005)	Delboni, Borges, Laender, and Jr (2007), Dolbear, Hart, and Goodwin (2007), Yao and Thill (2005)	Barclay and Galton (2013), Gapp (1995)	Barouni and Moulin (2015), Denofsky (1976), Ferenz (2000), Gapp (1995), Kelleher and Costello (2009), Wallgrün, Klippel, and Baldwin (2014), Xu and Klippel (2012), Yao and Thill (2005)
	PDF		Denofsky (1976)		Schockaert, De Cock, and Kerre (2008)		Hall and Jones (2012), Schockaert, De Cock, and Kerre (2008), Wallgrün, Klippel, and Baldwin (2014)	Derungs and Purves (2016), Dolbear, Hart, and Goodwin (2007)		Denofsky (1976), Ma, Kalashnikov, and Mehrotra (2008), Wallgrün, Klippel, and Baldwin (2014)
	Cost Function									
	Network Distance (flat)		Heath and Goodwin (2011)	Carbon (2007), Montello (1997)	Allen (1997)		Hall and Jones (2012)	Brooks, Kaufmann, and Lichtenstein (2008) Heath and Goodwin (2011)		Gryl et al. (2002)
	Network Distance (hierarchical)	Byrne (1979)	Marquet and Miralles-Guasch (2014)		Heath and Goodwin (2011), Kuipers (2000)	Couclelis, Golledge, Gale, and Tobler (1987)	Couclelis, Golledge, Gale, and Tobler (1987), Kuipers (2000), McNamara (1986)	Byrne (1979), Kuipers (2000), McNamara (1986)		Couclelis, Golledge, Gale, and Tobler (1987)
	Travel Time			Carbon (2007), Lee (1970)	Dao and Thill (2009), Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012), MacEachren (1980)		Minock and Mollevik (2013)	Hubbard (1978), MacEachren (1980), Marquet and Miralles-Guasch (2014)		Dao and Thill (2009)

Table 9: CFs associated with both spatial objects on a qualitative scale

Scale	Distance Measure	Barriers	Channels	Cultural	Distractor Object	Direction	Familiarity	Hierarchical Effects	Memory	Dynamic Elements
Nominal	Topological			Mark, Turk, and Stea (2007)				Bateman, Hois, Ross, and Tenbrink (2010), Richter, Winter, Richter, and Stirling (2013)		
	Mereotopological							Grütter, Helm-ing, Speich, and Bernstein (2011), Grütter, Scharrenbach, and Waldvogel (2010), Özçep, Grütter, and Möller (2012a, 2012b)		
Ordinal	Relative Distance (ordinal)	Gahegan (1995)	Coventry and Garrod (2004)		Gahegan (1995), Hund and Plumert (2007)	Frank (1996)	Hall and Jones (2012)			
	Fuzzy Membership							Robinson (2003)		
	Composition of Intervals					Frank (1992), Zwarts and Winter (2000)		Hernández, Clementini, and Di Felice (1995)		
Interval	Relative Distance (interval)	Carbon and Leder (2005)Sadalla, Burroughs, and Staplin (1980)				Sadalla, Burroughs, and Staplin (1980)		Hernandez (1994)	Sadalla, Burroughs, and Staplin (1980)	
	Non-Eucledian Geometry									
	Similarity		Schwering and Raubal (2005)							

Table 10: CFs associating with both spatial objects on a quantitative scale

Scale	Distance Measure	Barriers	Channels	Cultural	Distractor Object	Direction	Familiarity	Hierarchical Effects	Memory	Dynamic Elements
Ratio	Magnitude Estimation	Baird, Wagner, and Nomu (1982), MacEachren (1980), Qi, Shu, and Xu (2006), Wagner (2006)			Radvansky, Carlson-Radvansky, and Irwin (1995), Wagner (2006)	Lundberg and Ekman (1973)	Wagner (2006)	MacEachren (1980), Thorndyke (1981)	McNamara and Diwadkar (1997), Thorndyke (1981), Wagner (2006)	
	Ratio Estimation				Burigo and Coventry (2010)	Carlson and Covey (2005)				
	Absolute Distance	Gahegan (1995), Hart and Dolbear (2013), Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012), Yao and Thill (2005)	Brennan and Martin (2012), Hart and Dolbear (2013)		Carlson and Logan (2001), Denofsky (1976), Dolbear, Hart, and Goodwin (2007), Gahegan (1995), Kelleher and Costello (2009), Yao and Thill (2005)	Carlson and Kenny (2006), Kelleher and Kruijff (2005), Skoumas, Pfoser, and Kyrillidis (2013, 2014), Yao and Thill (2005)	Barouni and Moulin (2015), Brennan and Martin (2012), Gapp (1995), Hall and Jones (2012), Yao and Thill (2005)	Brennan and Martin (2012), Dolbear, Hart, and Goodwin (2007), Roller, Speriosu, Rallapalli, Wing, and Baldrige (2012)	Byrne (1979)	Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012)
	PDF	Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012), Yao and Thill (2005)			Denofsky (1976)	Skoumas, Pfoser, and Kyrillidis (2013, 2014)	Hall and Jones (2012)			
	Cost Function		Huriot, Smith, and Thisse (1989)							
	Network Distance (flat)	Gryl et al. (2002), Heath and Goodwin (2011), Montello (1997)	Gryl et al. (2002)		Montello (1997)	A. Stevens and Coupe (1978)	Dao and Thill (2009), Hall and Jones (2012), Montello (1997)	Brennan and Martin (2012), Heath and Goodwin (2011)		
	Network Distance (hierarchical)				McNamara, Hardy, and Hirtle (1989)	Couclelis, Golledge, Gale, and Tobler (1987)		Heath and Goodwin (2011), Kuipers (2000)		
	Travel Time	Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012), MacEachren (1980), Montello (1997)			Dao and Thill (2009), Hubbard (1978)	Lee (1970)	Dao and Thill (2009), Hourihan and Jones (1979)	Cai (2007)		Minock and Mollevik (2013)

Table 11: Nearness Tasks

Generalised Task	Type of Task	Question associated with "near"	Primary feature								CFs											Paper			
			Mereotopological	Relative Distance (ordinal)	Rel.Distance	Absolute Distance	PDF	Network Distance (flat)	Network Distance (hierarchical)	Travel Time	Barriers	Channels	Configuration	Density of Objects	Direction	Distribution of Objects	Dynamic Elements	Emotional Involvement	Familiarity	Hierarchical Effects	Population Density	Reachability	Size	Object Type	
Route Description	Provide directions Robot instruction	How to go to a cultural event? How to instruct robot for navigation?							X X												X X X	X X X	X X X		Kuipers (2000) Kuipers (2000)
Referential identification (Which?)	Tourism	Which interesting places are nearby each other?			X			X				X							X	X		X		X	Brennan and Martin (2012)
	Planning to open a bookshop	How many (which) bookshops are nearby?			X											X									Pareti and Klein (2011)
	Visit nearby recreation	Which are the nearby forest?			X	X				X	X						X					X			Kienast, Degenhardt, Weilenmann, Wäger, and Buchecker (2012)
	Shopping	Which shops are nearby for shopping?																		X					Cai (2007)
	Activity	Which objects are near for activity in general? (going to school, shopping, recreation, going to theater)				X					X		X	X				X	X		X	X		X	Yao and Thill (2005)
	Shopping	Which shops are near?								X				X							X				Hubbard (1978)
	Watch TV	Which couches are near for watching TV?				X							X										X	X	Ferenz (2000)
	Locate nearby objects	Which restaurants are nearby?				X				X												X			Hahn, Fogliaroni, Frank, and Navratil (2016)
Spatial locating	Which restaurants are nearby?															X							X	Minock and Mollevik (2013)	
Traveling: locate POIs, objects	Which objects are near a POI?				X																	X		X	Schockaert, De Cock, and Kerre (2008)
Spatial Description (Where?)	Sharing location on social networks	What objects are nearby users?			X																			X	Canneyt, Crommenlaan, Schockaert, Laere, and Dhoedt (2012)
	Traveling	Where is a petrol station nearby?			X	X					X			X								X		X	Gahegan (1995)
	Rescue mission	Where nearby is the accident happening?					X				X														Ma, Kalashnikov, and Mehrotra (2008)
	Identify objects from news	Where nearby is the event happening?				X								X										X	Xu and Klippel (2012)
	Proximity trip	Where is a person going for a nearby trip?							X					X								X			Marquet and Miralles-Guasch (2014)
	Short trip	Where are the nearby public services?	X																		X				Grütter, Scharrenbach, and Waldvogel (2010)
	Consuming public services	Where are the persons commuting to?	X																		X			X	Grütter, Helming, Speich, and Bernstein (2011)
	Daily commute	Where is a Postbox?	X																					X	Coventry and Garrod (2004)
Activity general	Assess the location of a fiber break, to dispatch crew (rescue crew)	X			X						X								X			X		X	Barouni and Moulin (2015)
Combination of spatial description and referential identification	Locating Spatial Objects	Which objects are located where nearby?			X	X													X				X		Skoumas, Pfoser, and Kyriallidis (2013, 2014)
	Map-Task	Which objects are where on a map located?			X												X								Cai, Wang, MacEachren, and Fuhrmann (2005)
	Organise Rescue Mission	Which robot can be used where nearby? (Multi-robot allocation problem)			X			X														X			Flushing, Gambardella, and Di Caro (2013)
	Land-Use Planning	Where should a waste dump be planted?			X		X				X									X	X		X	X	Brennan and Martin (2012)
	Land-Use Planning	Position Radio-tower as efficient as possible			X		X				X									X	X		X	X	Brennan and Martin (2012)

References

- Allen, G. L. (1997). From Knowledge to Words to Wayfinding : Issues in the Production and Comprehension of Route Directions. In Stephen C. Hirtle & A. U. Frank (Eds.), *Spatial Information Theory A Theoretical Basis for GIS. COSIT 1997* (pp. 363–372). LNCS vol. 1329. Berlin, Heidelberg, New York.
- Altman, D. (1994). Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal of Geographical Information Systems*, 8(3), 271–289.
- Ashley, A., & Carlson, L. A. (2007). Encoding direction when interpreting proximal terms. *Language and Cognitive Processes*, 22(7), 1021–1044.
- Baird, J. C., Wagner, M., & Nomu, E. (1982). Impossible Cognitive Spaces. *Geographical*, 14(3), 204–216.
- Barclay, M., & Galton, A. (2013). Selection of reference objects for locative expressions : the importance of knowledge and perception. In T. Tenbrink, J. Wiener, & C. Claramunt (Eds.), *Representing Space in Cognition*. Oxford: Oxford University Press.
- Barouni, F., & Moulin, B. (2015). An Intelligent Spatial Proximity System Using Neurofuzzy Classifiers and Contextual Information. *Geomatica*, 69(3), 285–296.
- Barsalou, L. W. (1985). Ideals, Central Tendency, and Frequency of Instantiation as Determinants of Graded Structure in Categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 629–654.
- Barwise, J., & Seligman, J. (1997). *Information Flow: The Logic of Distributed Systems*. Cambridge: Cambridge University Press.
- Bateman, J. (2010a). Language and Space: a two-level semantic approach based on principles of ontological engineering. *International Journal of Speech Technology*, 13(1), 29–48.
- Bateman, J. (2010b). Situating Spatial Language and the Role of Ontology: Issues and Outlook. *Language and Linguistics Compass*, 4(8), 639–664.
- Bateman, J., Hois, J., Ross, R., & Tenbrink, T. (2010). A linguistic ontology of space for natural language processing. *Artificial Intelligence*, 174(14), 1027–1071.
- Bateman, J., Tenbrink, T., & Farrar, S. (2007). The Role of Conceptual and Linguistic Ontologies in Interpreting Spatial Discourse. *Discourse Processes*, 44(3), 175–212.
- Bazire, M., Bazire, M., Brézillon, P., & Brézillon, P. (2005). Understanding Context Before Using It. In A. Dey, B. Kokinov, D. Leake, & R. Turner (Eds.), *Modeling and Using Context. CONTEXT 2005* (pp. 29–40). LNCS vol 3554. Berlin, Heidelberg: Springer-Verlag.
- Bera, R., & Claramunt, C. (2003). Topology-based proximities in spatial systems. *Journal of Geographical Systems*, 5(4), 353–379.
- Brennan, J. (2009). *A Framework for Modelling Spatial Proximity* (PhD Thesis, University of New South Wales).
- Brennan, J., & Martin, E. (2002). Foundations for a Formalism of Nearness. In R. I. McKay & J. Slaney (Eds.), *AI 2002: Advances in Artificial Intelligence. AI 2002* (pp. 71–82). LNCS vol 2557. Berlin, Heidelberg: Springer.
- Brennan, J., & Martin, E. (2006). Membership Functions for Spatial Proximity. In A. Sattar & B. Kang (Eds.), *AI 2006: Advances in Artificial Intelligence* (pp. 942–949). Lecture Notes in Artificial Intelligence vol 4304. Springer.
- Brennan, J., & Martin, E. (2012). Spatial proximity is more than just a distance measure. *International Journal of Human-Computer Studies*, 70(1), 88–106.
- Brennan, J., Martin, E., & Mihye, K. (2004). Developing An Ontology Of Spatial Relations. In J. S. Gero, B. Tversky, & T. Knight (Eds.), *Visual and Spatial Reasoning in Design III* (pp. 163–182). Sydney: Key Centre of Design Computing and Cognition, University of Sydney.

- Briggs, R. (1973). Urban Cognitive Distance. In R. M. Downs & D. Stea (Eds.), *Image and Environment* (pp. 361–388). Chicago: Aldine.
- Brooks, C. M., Kaufmann, P. J., & Lichtenstein, D. R. (2008). Trip chaining behavior in multi-destination shopping trips: A field experiment and laboratory replication. *Journal of Retailing*, 84(1), 29–38.
- Burigo, M., & Coventry, K. R. (2010). Context Affects Scale Selection for Proximity Terms. *Spatial Cognition & Computation*, 10(4), 292–312.
- Burigo, M., Coventry, K. R., & Cangelosi, A. (2016). Spatial Language and Converseness. *Quarterly Journal of Experimental Psychology*, 11, 1–19.
- Byrne, R. W. (1979). Memory for Urban Geography. *Quarterly Journal of Experimental Psychology*, 31(1), 147–154.
- Cai, G. (2007). Contextualization of Geospatial Database Semantics for Human–GIS Interaction. *Geoinformatica*, 11(2), 217–237.
- Cai, G. (2009). Modelling Contextual Knowledge for Adaptive Geographic Visualization. In H. Lin & M. Batty (Eds.), *Virtual Geographic Environment* (pp. 151–163). Beijing: Science Press.
- Cai, G., Wang, H., & MacEachren, A. M. (2003). Communicating Vague Spatial Concepts in Human-GIS Interactions: A Collaborative Dialogue Approach. In W. Kuhn, M. F. Worboys, & S. Timpf (Eds.), *Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003* (pp. 287–300). LNCS vol. 2825. Berlin, Heidelberg: Springer.
- Cai, G., Wang, H., MacEachren, A. M., & Fuhrmann, S. (2005). Natural Conversational Interfaces to Geospatial Databases. *Transactions in GIS*, 9(2), 199–221.
- Canneyt, S. V., Crommenlaan, G., Schockaert, S., Laere, O. V., & Dhoedt, B. (2012). Using Social Media to Find Places of Interest : A Case Study. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information*, ACM.
- Carbon, C.-C. (2007). Autobahn People: Distance Estimations Between German Cities Biased by Social Factors and the. In T. Barkowsky, M. Knauff, G. Ligozat, & D. R. Montello (Eds.), *Spatial Cognition V: Reasoning, Action, Interaction* (pp. 489–500). LNCS vol. 4387. Berlin Heidelberg: Springer.
- Carbon, C.-C., & Leder, H. (2005). The Wall inside the brain: Overestimation of distances crossing the former Iron Curtain. *Psychonomic Bulletin & Review*, 12(4), 746–750.
- Carlson, L. A., & Covey, E. S. (2005). How far Is "near"? Inferring distance from spatial descriptions. *Language and Cognitive Processes*, 20(5), 617–631.
- Carlson, L. A., & Kenny, R. (2006). Interpreting spatial terms involves simulating interactions. *Psychonomic Bulletin & Review*, 13(4), 682–688.
- Carlson, L. A., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, 29(6), 883–892.
- Carlson, L. A., & Van Deman, S. R. (2004). The space in spatial language. *Journal of Memory and Language*, 51(3), 418–436.
- Cohn, A. G., & Renz, J. (2008). Qualitative Spatial Representation and Reasoning. In F. van Harmelen, V. Lifschitz, & B. Porter (Eds.), *Handbook of Knowledge Representation* (Chap. 13, pp. 551–596). Amsterdam, Oxford: Elsevier B.V.
- Costello, F. J., & Kelleher, J. D. (2006). Spatial Prepositions in Context: The Semantics of near in the Presence of Distractor Objects. In *Third ACL-SIGSEM Workshop on Prepositions. Preposition '06*.
- Couclelis, H., Gollledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology*, 7(2), 99–122.
- Coventry, K. R. (2013). On the Mapping between Spatial Language and the Vision and Action Systems. In Y. Coello & A. Bartolo (Eds.), *Language and Action in Cognitive Neuroscience* (Chap. Chapter 11, pp. 209–225). East Sussex, UK: Psychology Press.

- Coventry, K. R., & Garrod, S. C. (2004). *Saying, Seeing, and Acting: the Psychological Semantics of Spatial Prepositions*. Essays in Cognitive Psychology. Hove: Psychology Press.
- Crangle, C., & Suppes, P. (1989). Geometrical Semantics for Spatial Prepositions. *Midwest Studies in Philosophy*, 14, 399–422.
- Crangle, C., & Suppes, P. (1994). *Language and Learning for Robots*. Stanford: CSLI Publications.
- Dale, R., & Reiter, E. (1995). The Role of the Gricean Maxims in the Generation of Referring Expressions. *Cognitive Science*, (18), 233–263.
- Dao, T. H. D., & Thill, J. C. (2009). Contextualized Space-Time Accessibility Modeling and Measurements with NeuroFuzzy Proximity Relations. In *Proceedings of the International Conference on Advanced Geographic Information Systems and Web Services*. GEOWS 2009 (pp. 158–164). IEEE Computer Society.
- Dehane, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. Oxford, New York: Oxford University Press.
- Delboni, T. M., Borges, K. A. V., Laender, A. H. F., & Jr, C. A. D. (2007). Semantic Expansion of Geographic Web Queries Based on Natural Language Positioning Expressions. *Transactions in GIS*, 11(3), 377–397.
- Denofsky, M. E. (1976). *How near is near? a near specialist*. AI Memo No. 344. Cambridge, MA: Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
- Derungs, C., & Purves, R. S. (2016). Mining nearness relations from an n-grams Web corpus in geographical space. *Spatial Cognition & Computation*, 16(4), 301–322.
- Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, 5(1), 4–7.
- Dolbear, C., Hart, G., & Goodwin, J. (2007). From Theory to Query: Using ontologies to make explicit imprecise spatial relationships for database querying. In *Conference on Spatial Information Theory (COSIT 2007)*.
- Downs, R. M., & Stea, D. (1973). Cognitive Maps and Spatial Behaviour : Process and Products. In R. M. Downs & D. Stea (Eds.), *Image and Environment: Cognitive Mapping and Spatial Behavior*. New Brunswick: Aldine Transactions.
- Egenhofer, M. J., & Mark, D. M. (1995). Naive Geography. In *Spatial Information Theory A Theoretical Basis for GIS*. COSIT 1995 (September, pp. 1–15). LNCS vol. 988.
- Ekman, G., & Bratfisch, O. (1965). Subjective distance and emotional involvement. A psychological mechanism. *Acta psychologica*, 24(5), 430–437.
- Ferenz, K. (2000). *The role of nongeometric information in spatial language*. (Unpublished doctoral dissertation, Dartmouth College, Hanover, NH).
- Fisher, P. (2000). Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113(1), 7–18.
- Fisher, P. F., & Orf, T. M. (1991). An investigation of the meaning of near and close on a university campus. *Comput. Environ. and Urban Systems*, 15, 23–35.
- Flushing, E. F., Gambardella, L. M., & Di Caro, G. A. (2013). Strategic Control of Proximity Relationships in Heterogeneous Search and Rescue Teams. In *Proceedings of the 3rd IROS Workshop on Robots and Sensors integration in future rescue INformation system*. ROSIN, Tokyo.
- Frank, A. U. (1992). Qualitative spatial reasoning about distances and directions in geographic space. *Journal of Visual Languages & Computing*, 3(4), 343–371.
- Frank, A. U. (1996). Qualitative spatial reasoning: cardinal directions as an example. *International journal of geographical information systems*, 10(3), 269–290.
- Frank, A. U. (2006). Twenty Years of Reasoning with Spatial Relations. In P. Fischer (Ed.), *Classics from IJGIS: Twenty years of the International Journal of Geographical Information Science and Systems* (pp. 353–361). CRC Press.

- Freundschuh, S. M. (1998). Cognitive distance at various geographic scales. In M. J. Egenhofer & R. Golledge (Eds.), *Spatial and Temporal Reasoning in Geographic Information Systems* (pp. 212–34). Oxford: Oxford University Press.
- Gahegan, M. (1995). Proximity operators for qualitative spatial reasoning. In A. U. Frank & W. Kuhn (Eds.), *Spatial Information Theory A Theoretical Basis for GIS*. COSIT 1995 (pp. 31–44). LNCS vol. 4128. Berlin, Heidelberg: Springer.
- Galton, A. (2009). Spatial and temporal knowledge representation. *Earth Science Informatics*, 2(3), 169–187.
- Gapp, K.-P. (1995). Object Localization: Selection of Optimal Reference Objects. In A. U. Frank & W. Kuhn (Eds.), *Spatial Information Theory - A Theoretical Basis for GIS*, Berlin, Heidelberg, New York: Springer.
- Golledge, R. G., & Zannaras, G. (1973). Cognitive approaches to the analysis of human spatial behaviour. In W. Ittelton (Ed.), *Environmental Cognition* (pp. 59–94). New York: Academic Press.
- Grice, H. (1975). Logic and Conversation. In *Speech Acts* (Vol. 3, pp. 43–58). Syntax and Semantics. New York: Academic Press.
- Grütter, R., Helming, I., Speich, S., & Bernstein, A. (2011). Rewriting Queries for Web Searches That Use Local Expressions. In *5th International Symposium on Rules. RuleML 2011* (pp. 345–359). Barcelona.
- Grütter, R., Scharrenbach, T., & Waldvogel, B. (2010). Vague Spatio-Thematic Query Processing: A Qualitative Approach to Spatial Closeness. *Transactions in GIS*, 14(2), 97–109.
- Gryl, A., Moulin, B., Kettani, D., Gryll, A., Moulin, B., & Kettani, D. (2002). A Conceptual Model for Representing Verbal Expressions used in Route Descriptions. *Spatial Language*, 19–42.
- Guesgen, H. W. (2002). Reasoning About Distance Based on Fuzzy Sets. *Applied Intelligence*, 17, 265–270.
- Guesgen, H. W., & Albrecht, J. A. (2000). Imprecise reasoning in geographic information systems. *Fuzzy Sets and Systems*, 113(1), 121–131.
- Gurevych, I., Porzel, R., & Malaka, R. (2006). In Context: Integrating Domain- and Situation-specific Knowledge. In W. Wahlster (Ed.), *SmartKom - Foundations of Multimodal Dialogue Systems* (pp. 269–284). Berlin, Heidelberg: Springer.
- Hahn, J., Fogliarini, P., Frank, A. U., & Navratil, G. (2016). A Computational Model for Context and Spatial Concepts. In T. Sarjakoski, Y. S. Maribel, & L. T. Sarjakoski (Eds.), *Geospatial Data in a Changing World: Selected Papers of the 19th AGILE Conference on Geographic Information Science* (pp. 3–19). Lecture Notes in Geoinformation and Cartography. Springer.
- Hall, M. M., & Jones, C. B. (2012). Cultural and Language Influences on the Interpretation of Spatial Prepositions. In *GI_Forum 2012: Geovisualization, Society & Learning*, Berlin/Offenbach.
- Hand, D. J. (1996). Statistics and the Theory of Measurement. *Journal of the Royal Statistical Society. Series A*(159), 445–492.
- Hart, G., & Dolbear, C. (2013). *Linked Data: A Geographic Perspective*. Boca Raton: CRC Press.
- Hawes, N., Klenk, M., Lockwood, K., Horn, G. S., & Kelleher, J. D. (2012). Towards a Cognitive System that Can Recognize Spatial Regions Based on Context. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence* (pp. 200–206). AAAI Publications.
- Heath, T., & Goodwin, J. (2011). Linking Government Data for Government and Consumer Applications. In D. Wood (Ed.), *Linking Government Data* (pp. 73–92). New York, NY: Springer New York.
- Hedblom, M. M., Kutz, O., & Neuhaus, F. (2015). Image Schemas as Families of Theories. In *Proceedings of the 4th International Workshop, C3GI* (pp. 19–33). Istanbul.

- Helming, I., Bernstein, A., Grütter, R., & Vock, S. (2011). Making close to Suitable for Web Search A Comparison of Two Approaches. In *Terra Cognita - Foundations, Technologies and Applications of the Geospatial Web*, Bonn.
- Hernández, D. (1994). *Qualitative Representation of Spatial Knowledge*. Berlin, Heidelberg, New York: Springer-Verlag Heidelberg.
- Hernández, D., Clementini, E., & Di Felice, P. (1995). Qualitative Distances. In *Spatial Information Theory A Theoretical Basis for GIS. COSIT 1995* (pp. 45–57). LNCS vol. 988. Berlin, Heidelberg: Springer.
- Herskovits, A. (1986). *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English*. Studies in Natural Language Processing. Cambridge: Cambridge University Press.
- Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 3(13), 208–217.
- Hourihan, K. J., & Jones, L. E. (1979). Distance Estimation and Cognitive Mapping: A Multidimensional Scaling Analysis. *The Irish Journal of Psychology*, 4(2), 79–94.
- Hubbard, R. (1978). A Review of Selected Factors Conditioning Consumer Travel Behavior. *Journal of Consumer Research*, 5(1), 1.
- Hund, A. M., & Plumert, J. M. (2007). What Counts as By? Young Children's Use of Relative Distance to Judge Nearby. *Developmental Psychology*, 43(1), 121–133.
- Huriot, J.-M., Smith, T. E., & Thisse, J.-F. (1989). Minimum-Cost Distances in Spatial Analysis. *Geographical Analysis*, 21(4), 294–315.
- Inge, C. A., & Ghosh, A. (1990). Consumer and Produced Behaviour in a Multipurpose Shopping Environment. *Geographical Analysis*, 22(1), 70–93.
- Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, 52(2), 93–129.
- Johnson, M. (1987). *The Body in the Mind*. Chicago, London: The University of Chicago Press.
- Kelleher, J. D., & Costello, F. J. (2009). Applying computational models of spatial prepositions to visually situated dialog. *Computational Linguistics*, 35(2), 271–306.
- Kelleher, J. D., & Kruijff, G.-J. M. (2005). A context-dependent algorithm for generating locative expressions in physically situated environments. In G. Wilcock, K. Jokinen, C. Mellish, & E. Reiter (Eds.), *Proceedings of the 10th European Workshop on Natural Language Generation. ENLG* (pp. 68–74). Aberdeen.
- Kelleher, J. D., Kruijff, G.-J. M., & Costello, F. J. (2006). Proximity in context: an empirically grounded computational model of proximity for processing topological spatial expressions. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL* (pp. 745–752). Association for Computational Linguistics.
- Kemmerer, D. (1999). 'Near' and 'far' in language and perception. *Cognition*, 73(1), 35–63.
- Kettani, D., & Moulin, B. [Bernard]. (1999). A Spatial Model Based on the Notions of Spatial Conceptual Map and of Object's Influence Areas. In C. Freska & D. M. Mark (Eds.), *COSIT'99* (pp. 401–416). LNCS vol. 988. Berlin, Heidelberg: Springer.
- Khan, A., Vasardani, M., & Winter, S. (2013). Extracting Spatial Information From Place Descriptions. In *Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place (COMP 2013)* (pp. 62–69). New York.
- Kienast, F., Degenhardt, B., Weilenmann, B., Wäger, Y., & Buchecker, M. (2012). GIS-assisted mapping of landscape suitability for nearby recreation. *Landscape and Urban Planning*, 105(4), 385–399.
- Kracht, M. (2008). *Language and Space*, Department of Linguistics, UCLA, Los Angeles.

- Kruijff, G.-J. M., Lison, P., Benjamin, T., Jacobsson, H., & Hawes, N. A. (2007). Incremental, multi-level processing for comprehending situated dialogue in human-robot interaction. In L. Lopes, T. Belpaeme, & S. Cowley (Eds.), *Symposium on Language and Robots. LangRo 2007* (pp. 1–36). Aveiro.
- Kuipers, B. (1978). Modeling Spatial Knowledge. *Cognitive Science*, 2(2), 129–153.
- Kuipers, B. (2000). The Spatial Semantic Hierarchy. *Artificial Intelligence*, 119(1), 191–233.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- LeBoeuf, R., & Shafir, E. (2009). Anchoring on the “Here” and “Now” in Time and Distance Judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 81–93.
- Lee, T. (1970). Perceived Distance as a Function of Direction in the City. *Environment and Behavior*, 2(1), 40–51.
- Levinson, S. C. (1996). Language and Space. *Annual Review of Anthropology*, 25(1), 353–382.
- Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: The MIT Press.
- Liu, Y., Guo, Q. H., Wieczorek, J., & Goodchild, M. F. (2009). Positioning localities based on spatial assertions. *International Journal of Geographical Information Science*, 23(11), 1471–1501.
- Lloyd, R., & Heivlyt, C. (1987). Systematic Distortions in Urban Cognitive Maps. *Annals of the Association of American Geographers*, 77(2), 191–207.
- Logan, G. D., & Sadler, D. D. (1996). A Computational Analysis of the Apprehension of Spatial Relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and Space* (Chap. 13, pp. 493–529). Cambridge, MA, USA: The MIT Press.
- Lu, B., Charlton, M., Harris, P., & Stewart, A. (2014). Geographically weighted regression with a non-Euclidean distance metric : a case study using hedonic house price data. *International Journal of Geographical Information Science*, 28(4), 660–681.
- Luce, R. D. (1956). Semiorders and a Theory of Utility Discrimination. *Econometrica*, 24(2), 178–191.
- Luce, R. D., Krantz, D. H., Suppes, P., & Tversky, A. (1990). *Foundations of Measurement Volume III: Representation, Axiomatization, and Invariance*. San Diego, London: Academic Press.
- Lundberg, U., & Ekman, G. (1973). Subjective geographic distance: A multidimensional comparison. *Psychometrika*, 38(May 1971), 113–122.
- Lynch, K. (1960). *The Image of the City*. Cambridge, MA: MIT Press.
- Ma, Y., Kalashnikov, D. V., & Mehrotra, S. (2008). Towards Managing Uncertain Spatial Information for Situational Awareness Applications. *IEE Transactions on Knowledge and Data Engineering*, 20(10).
- MacEachren, A. M. (1980). Travel Time As the Basis of Cognitive Distance. *The Professional Geographer*, 32(1), 30–36.
- Mainwaring, S. D., Tversky, B., Ohgishi, M., & Schiano, D. J. (2003). Descriptions of Simple Spatial Scenes in English and Japanese. *Spatial Cognition & Computation*, 3(1), 3–42.
- Maki, R. H. (1981). Categorization and Distance Effects With Spatial Linear Orders. *Journal of Experimental Psychology*, 7(1), 15–22.
- Mark, D. M., & Egenhofer, M. J. (1994). Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subjects Testing. *Cartography and Geographic Information Systems*, 21(4), 195–212.
- Mark, D. M., Turk, A. G., & Stea, D. (2007). Progress on Yindjibarndi Ethnophysiography. In S. Winter, M. Duckham, L. Kulik, & B. Kuipers (Eds.), *Spatial Information Theory. COSIT 2007* (pp. 1–19). LNCS vol. 4736. Berlin Heidelberg: Springer.

- Marquet, O., & Miralles-Guasch, C. (2014). The Walkable city and the importance of the proximity environments for Barcelona's everyday mobility. *Cities*, 42, 258–266.
- McNamara, T. P. (1986). Mental Representations of Spatial Relations. *Cognitive Psychology*, 18(1), 87–121.
- McNamara, T. P., & Diwadkar, V. A. (1997). Symmetry and Asymmetry of Human Spatial Memory. *Cognitive Psychology*, 34(2), 160–90.
- McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective Hierarchies in Spatial Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 211–227.
- McNamara, T. P., Rump, B., & Werner, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin & Review*, 10(3), 589–595.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception*. Cambridge: Cambridge University Press.
- Miller, H. J., & Wentz, E. A. (2003). Representation and Spatial Analysis in Geographic Information Systems. *Annals of the Association of American Geographers*, 93(3), 574–594.
- Minock, M., & Mollevik, J. (2013). Context-dependent near and far in spatial databases via supervaluation. *Data & Knowledge Engineering*, 86(270019), 295–305.
- Montello, D. R. (1991). The measurement of cognitive distance: Methods and construct validity. *Journal of Environmental Psychology*, 11, 101–122.
- Montello, D. R. (1993). Scale and Multiple Psychologies of Space Daniel. In A. U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS. COSIT 1993* (pp. 312–321). LNCS vol. 716. Berlin, Heidelberg: Springer.
- Montello, D. R. (1997). The Perception and Cognition of Environmental Distance: Direct Sources of Information. In S. Hirtle & A. U. Frank (Eds.), *Spatial Information Theory A Theoretical Basis for GIS. COSIT 1997* (pp. 297–311). LNCS vol. 1329. Berlin: Springer.
- Morrow, D. G., & Clark, H. H. (1988). Interpreting Words in Spatial Descriptions. *Language and Cognitive Processes*, 3(4), 275–291.
- Nam, S. (1995). *The Semantics of Locative Prepositional Phrases in English* (PhD-Thesis, University of California, Los Angeles).
- Nasar, J. L., Valencia, H., Omar, Z. A., Chueh, S.-C., & Hwang, J.-H. (1985). Out of Sight Further from Mind: Destination Visibility and Distance Perception. *Environment & Behavior*, 17(5), 627–639.
- O'Keefe, J. (1996). The Spatial Prepositions in English , Vector Grammar , and the Cognitive Map Theory. In P. Bloom, M. A. Peterson, L. Nadel, & F. Merrill Garrett (Eds.), *Language and Space*. Cambridge, MA: MIT Press.
- Özçep, Ö. L., Grütter, R., & Möller, R. (2012a). Dynamics of a Nearness Relation — First Results. In M. Bhatt, H. W. Guesgen, & E. Davis (Eds.), *Proceedings of the International Workshop on Spatio-Temporal Dynamics (STeDy 2012)*.
- Özçep, Ö. L., Grütter, R., & Möller, R. (2012b). Nearness Rules and Scaled Proximity. In L. D. Raedt, C. Bessiere, D. Dubois, P. Doherty, P. Frasconi, F. Heintz, & P. Lucas (Eds.), *ECAI 2012, 20th European Conference on Artificial Intelligence* (pp. 636–641). Frontiers in Artificial Intelligence and Applications. Amsterdam, Berlin, Oxford, Tokyo, Washington, D.C: IOS Press.
- Pareti, P., & Klein, E. (2011). Learning Vague Concepts for the Semantic Web. In *Proceedings of the Joint Workshop on Knowledge Evolution and Ontology Dynamics. In Conjunction with ISWC 2011*. CEUR workshop proceedings. vol. 784.
- Pustejovsky, J., Moszkowicz, J., & Verhagen, M. (2012). A Linguistically Grounded Annotation Language for Spatial Information. *TAL Traitement Automatique des Langues*, 53(2), 87–113.

- Qi, C., Shu, H., & Xu, A. (2006). Formal Properties of Cognitive Distance in Geographical Space. In *Proceedings of the 16th International Conference on Artificial Reality and Telexistence-Workshops (ICAT'06)*.
- Radvansky, G., Carlson-Radvansky, L., & Irwin, D. (1995). Uncertainty in estimating distances from memory. *Memory & cognition*, 23(5), 596–606.
- Randell, D. A., Cui, Z., & Cohn, A. G. (1992). A spatial logic based on regions and connection. In *3rd Int. Conf. on Knowledge Representation and Reasoning* (Vol. 92, pp. 165–176). Morgan Kaufmann.
- Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, 130(2), 273–298.
- Richter, D., Winter, S., Richter, K.-F., & Stirling, L. (2013). Granularity of locations referred to by place descriptions. *Computers, Environment and Urban Systems*, 41, 88–99.
- Roberts, F. S. (1973). Tolerance Geometry. *Notre Dame Journal of Formal Logic*, 14(1), 68–76.
- Robinson, V. B. (2003). A Perspective on the Fundamentals of Fuzzy Sets and their Use in Geographic Information Systems. *Transactions in GIS*, 7(1), 3–30.
- Rodríguez, M. A., & Egenhofer, M. J. (2004). Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. *International Journal of Geographical Information Science*, 18(3), 229–256.
- Roller, S., Speriosu, M., Rallapalli, S., Wing, B., & Baldridge, J. (2012). Supervised Text-based Geolocation Using Language Models on an Adaptive Grid. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Association for Computational Linguistics.
- Sadalla, E. K., Burroughs, J. W., & Staplin, L. J. (1980). Reference Points in Spatial Cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 6(5), 516–528.
- Sanderson, M., & Kohler, J. (2004). Analyzing geographic queries. In *SIGIR Workshop on Geographic Information Retrieval* (Vol. 2).
- Schockaert, S., De Cock, M., & Kerre, E. E. (2008). Location approximation for local search services using natural language hints. *International Journal of Geographical Information Science*, 22(3), 315–336.
- Schwering, A., & Raubal, M. (2005). Spatial Relations for Semantic Similarity Measurement. In *Perspectives in Conceptual Modeling* (pp. 1–11).
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3), 379–423.
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: The University of Illinois Press.
- Sharma, J., Flewelling, D., & Egenhofer, M. J. (1994). A Qualitative Spatial Reasoner. In T. C. Waugh & R. G. Healey (Eds.), *Proceedings of the Sixth International Symposium on Spatial Data Handling* (pp. 665–681). London: Taylor & Francis.
- Skoumas, G., Pfoser, D., & Kyrillidis, A. (2013). On quantifying qualitative geospatial data: a probabilistic approach. In *Proceedings of the Second ACM International Workshop on Crowdsourced and Volunteered Geographic Information* (pp. 71–78).
- Skoumas, G., Pfoser, D., & Kyrillidis, A. (2014). *Location Estimation Using Crowdsourced Geospatial Narratives*. arXiv: 1408.5894
- Stevens, A., & Coupe, P. (1978). Distortions in Judged Spatial Relations. *Cognitive psychology*, 10(4), 422–437.
- Stevens, S. S., & Galanter, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 54(6), 377–411.
- Strzalecki, A. (1978). The relation between subjective distance and emotional involvement: Further experiment. *Acta Psychologica*, 42(5), 429–440.

- Suppes, P. (1991). The Principle of Invariance with Special Reference to Perception. In J. Doignon & J. Falmagne (Eds.), *Mathematical Psychology: Current Development* (pp. 35–53). New York: Springer-Verlag.
- Suppes, P., Krantz, D. H., Luce, R. D., & Tversky, A. (1989). *Foundations of Measurement Volume II: Geometrical, Threshold, and Probabilistic Representations*. San Diego, CA: Academic Press.
- Talmy, L. (2000). How language structures space. In *Toward a Cognitive Semantics* (Chap. 3, Vol. 1, pp. 177–254). Cambridge, MA: MIT Press.
- Tamminen, S., Oulasvirta, A., Toiskallio, K., & Kankainen, A. (2004). Understanding mobile contexts. *Personal and Ubiquitous Computing*, 8(2), 135–143.
- Taylor, H. A., & Tversky, B. (1996). Perspective in Spatial Descriptions. *Journal of Memory and Language*, 35(3), 371–391.
- Tezuka, T., Kambayashi, Y., & Takakura, H. (2001). Web-Based Inference Rules for Processing Conceptual Geographical Relationships. *Proceedings of the 2nd Int. Conf. on Web Information Systems Engineering, The 1st Int. Workshop on Web Geographical Information Systems*, 2, 14–21.
- Tezuka, T., Takakura, H., Lee, R., & Kambayashi, Y. (2001). Models for Conceptual Geographical Prepositions Based on Web Resource. *Journal of Geographic Information and Decision Analysis*, 5(2), 83–94.
- Thorndyke, P. W. (1981). Distance Estimation from Cognitive Maps. *Cognitive Psychology*, 13(4), 526–550.
- Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46, 234–240.
- Turney, P. D. (1996a). The Identification of Context-Sensitive Features: A Formal Definition of Context for Concept Learning. In *13th International Conference on Machine Learning (ICML96), Workshop on Learning in Context-Sensitive Domains* (pp. 53–59). Bari.
- Turney, P. D. (1996b). The Management of Context-Sensitive Features: A Review of Strategies. In *13th International Conference on Machine Learning (ICML96), Workshop on Learning in Context-Sensitive Domains* (pp. 60–66). Bari, Italy.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.)* 185(4157), 1124–31.
- Tversky, B. (1992). Distortions in Cognitive Maps. *Geoforum*, 23(2), 131–138.
- Tversky, B. (2000). Levels and Structure of Spatial Knowledge. In S. Freundschuh & R. Kitchen (Eds.), *Cognitive Mapping: Past, Present, and Future* (Chap. Chapter 3, pp. 24–44). Frontiers of Cognitive Science. London, New York: Routledge.
- Tversky, B. (2003). Structures Of Mental Spaces: How People Think About Space. *Environment & Behavior*, 35(1), 66–80.
- Tversky, B., Taylor, H. A., & Mainwaring, S. (1997). Langage et perspective spatiale. In M. Denis (Ed.), *Langage et cognition spatiale* (pp. 25–50). Science Cognitives. Paris: Masson.
- van der Zee, E., Adams, K., & Niemi, J. (2009). The Influence of Geometrical and Nongeometrical Features on the Use of the Lexical Concepts NEAR and FAR in English and Finnish. *Spatial Cognition and Computation*, 9(4), 305–317.
- van Benthem, J. F. A. K. (1991). *The Logic of Time: A Model-Theoretic Investigation into the Varieties of Temporal Ontology and Temporal Discourse* (Second Edi). Dordrecht: Springer.
- van Rooij, R. (2011). Vagueness , Tolerance and Non-Transitive Entailment. In P. Cintula, C. Fermüller, L. Godo, & P. Hajek (Eds.), *Understanding Vagueness - Logical, Philosophical and Linguistic Perspectives* (pp. 205–221). Studies in Logic vol. 36. London: College Publications.
- Vandeloise, C. (1991). *Spatial Prepositions: A Case Study from French*. Chicago: The University of Chicago Press.
- Vasardani, M., Winter, S., & Richter, K.-F. (2013). Locating place names from place descriptions. *International Journal of Geographical Information Science*, 27(12), 2509–2532.

- Vieu, L. (1991). *Sémantique des relations spatiales et inférences spatio-temporelles : Une contribution à l'étude des structures formelles de l'espace en Langage Naturel* (Doctoral dissertation, Université Paul Sabatier).
- Wagner, M. (1985). The metric of visual space. *Perception & Psychophysics*, 38(6), 483–495.
- Wagner, M. (2006). *The Geometries of Visual Space*. New Jersey: Lawrence Erlbaum Associates.
- Wallgrün, J. O., Klippel, A., & Baldwin, T. (2014). Building a Corpus of Spatial Relational Expressions Extracted from Web Documents. In *SIGSPATIAL '14*, Dallas/Fort Worth, TX, USA: ACM.
- Whittemore, R., & Knafl, K. (2005). The integrative review: updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553.
- Worboys, M. F. (1996). Metrics and topologies for geographic space. In *Advances in Geographic Information Systems Research II: Proceedings of the International Symposium on Spatial Data Handling* (Vol. 7), Delft, Netherlands.
- Worboys, M. F. (2001). Nearness relations in environmental space. *International Journal of Geographical Information Science*, 15(7), 633–651.
- Worboys, M. F. (2003). Communicating Geographic Information in Context. In M. Duckham, M. F. Goodchild, & M. F. Worboys (Eds.), *Foundations of Geographic Information Science* (pp. 33–45). London, New York: Taylor & Francis.
- Xiao, D., & Liu, Y. (2007). Study of Cultural Impacts on Location Judgments in Eastern China. *Spatial Information Theory*, 4736, 20–31.
- Xu, S., & Klippel, A. (2012). Developing Nearness Models from Geocoding Spatial Entities in a News Corpus. *giscience.org*, 1–4.
- Xu, S., & Klippel, A. (2013). Linking Context and Proximity Through Web Corpus. In *Proceedings of the 7th Workshop on Geographic Information Retrieval - GIR '13* (pp. 45–46). New York, New York, USA: ACM Press.
- Yao, X., & Thill, J.-C. (2005). How Far Is Too Far? - A Statistical Approach to Context-contingent Proximity Modeling. *Transactions in GIS*, 9(2), 157–178.
- Zwarts, J. (2017). Spatial semantics : Modeling the meaning of prepositions. *Language and Linguistics Compass*, 11(e12241), 1–20.
- Zwarts, J., & Gärdenfors, P. (2015). Locative and Directional Prepositions in Conceptual Spaces: The Role of Polar Convexity. *Journal of Logic, Language and Information*, 25(1), 109–138.
- Zwarts, J., & Winter, Y. (2000). Vector Space Semantics : A Model-Theoretic Analysis of Locative Prepositions. (1997), 169–211.