Making *close to* Suitable for Web Search
A Comparison of Two Approaches

Iris Helming[1], Abraham Bernstein[2], Rolf Grütter[1], and Stephan Vock[3]

[1] Swiss Federal Research Institute WSL, Birmensdorf, Switzerland {iris.helming, rolf.gruetter}@wsl.ch
[2] University of Zurich, Department of Informatics Zurich, Switzerland bernstein@ifi.uzh.ch
[3] stephan.vock@gmail.com

Abstract. In this paper we compare two approaches to model the vague german spatial relation *in der Nähe von* (English: "close to") to enable its usage in (semantic) web searches. A user wants, for example, to find all relevant documents regarding parks or forestal landscapes *close to* a city. The problem is that there are no clear metric distance limits for possibly matching places because they are only restricted via the vague natural language expression. And since human perception does not work only in distances we can't handle the queries simply with metric distances. Our first approach models the meaning of these expressions in description logics using relations of the Region Connection Calculus. A formalism has been developed to find all instances that are potentially perceived as *close to*. The second approach deals with the idea that everything that can be reached in a reasonable amount of time with a given means of transport (e.g. car) is potentially perceived as *close*. This approach uses route calculations with a route planner. The first approach has already been evaluated. The second is still under development. But we can already show a correlation between what people consider as *close to* and time needed to get there.
Keywords: Vague Spatial Relation, Local Expression, Region Connection Calculus (RCC), Route Planning, Reachability.

1 Introduction

Sometimes we want to search for places on the web and restrict the search results to a specific area. But we don't have an exact distance restriction in mind, we just want to look for something that is *close* or *not close to*, *a bit off* and so on. How can we make this restriction understandable to a search engine? So that future users could simply apply these expressions as keywords without further thinking about "translating" them?

Google[4] already delivers results for queries that include *near*. But these results show that it's not really taken care of the meaning of the preposition: if

[4] http://www.google.ch

you are looking for a "hotel in Zurich" for example, it returns also hotels which are in the area around Zurich. On the other hand, if disliking living in a city centre but wanting to get there quickly, you could search for "hotels near Zurich". The result will show you hotels near Zurich but also some which are located in the city centre. Also, these mechanisms can't scale with regards to the reference place (i.e. the place to which the first one is supposed to be near). So the area for searching doesn't have a bigger size if the reference place is bigger. With our knowledge representation approach (cf. [2]) such scaling is performed through the type of the administrative region of the reference place — the granularity of found nearby places is decreasing if the referent is situated on a more fine-grained scale, such as a village, or increasing if the referent is on a larger scale, such as a city. In this paper we will present a novel approach for decoding "nearness", which deals with statistical methods. We compare it to our previous knowledge engineering approach. With the new approach scaling works through the chosen means of transport — things that are near while driving a car may not be near while walking. Using this approach one could, in a future implementation, present the user the best results of *nearby* places for his traveling context.

Both approaches are meant to map the vague concepts of spatial relations that occur in natural language onto concrete geographical regions or places.

2 Related Work

Schokaert, De Cock and Kerre [8] suggest augmenting the structured information available to a local search service, such as Google Maps, with information extracted from the web. They show how nearness information in natural language and information about the surrounding neighborhood of a place can be translated into fuzzy restrictions and how such fuzzy restrictions can be used to estimate the location of a place with an unknown address. The vast amount of data addressed by the authors, together with the kinds of examples they provide, suggest that their approach is targeted on mass searches. In our case, the resources on the web, which could possibly be used to augment the searches, are sometimes scarce. Our second approach also is a statistical one, but we are using context information -traveling time- to give the best matching results for a special purpose.

Yao and Thill [9] also follow a statistical approach to handle vague natural language expressions of distance. Different to us, they directly transform their results to distances. We are avoiding this since human perception of vague spatial expressions doesn't work on metrical distances.

Also their discussion of general problems when dealing with vague expressions for distances is of interest for our approach. They are highlighting the importance of the context when a person has to judge if a place is near another place. Among others they name transport mode as an influence factor for the perception of nearness, e.g. Is city A close to city B? Yes for airplane, no for car. With the statistical approach we make use of this influence factor and show how contextual information in terms of means of transport can be modelled.

Mata [5] presents an approach to geographic information retrieval integrating topological, geographical and conceptual matching. For topological matching topological relations are extracted from overlaying data layers; for geographical matching constraints are obtained from dictionaries; for conceptual matching a geographic ontology is used. A constraint defines two geographic objects (points or polygons) as near provided they are connected by a third object (an arc, e.g., a road), the length of which is less than a given distance. Different from the approaches we are comparing, a metric distance measure thus is a necessary condition for nearness, although not a sufficient. However, the framework seems general enough to be aligned with that presented in section 3.

3 Knowledge Representation Approach

In this section we briefly summarize the important aspects of the knowledge representation approach, which we presented in [2]. For modelling nearness with this first approach we use information of the administrative structure of Switzerland, which can be obtained easily. It is freely available as a download for many countries. An administrative region like a district is responsible for administrative tasks like providing schools, medical supply, organizing elections and so on. Often borders for such regions are grown where also natural barriers like big streams exist. It has been shown before that the partitioning of a country into smaller parts has influence on human perception of space. Maki [3] for example showed that the affiliation to a state plays an important role in human perception of locations. In an experiment, subjects had to decide about the location of two cities regarding their orientation east-west. If the cities in question belong to different states, the reaction times were significantly shorter than with cities which belong to the same state. Human beings are able to judge faster about entities on a continuum if they can make use of category information. Among others, Carbon and Leder [1] showed that the membership to different political systems, structures or hierarchies influences the estimation of distance between two cities. They used estimation tasks for distances between cities east and west of the former border inside of Germany. Compared to pairs inside the same part of the former republic, distances were overestimated if the cities in question belonged to different parts.

Topologies of regions can be relatively easy obtained via modern geographic information systems (GIS) or spatial databases. Types of administrative regions and toplogical relations between them provide us with the possibility to reason on these regions as polygons. Randell, Cui and Cohn [7] developed a a set of spatial relations in first-order logic, the Region Connection Calculus (RCC), which we use for it. How this works is described in the next section. With this approach we provide a qualitative method for qualitative search queries. As Mark and Egenhofer already conclude metric is not the most important parameter of the semantics of most spatial natural-language expressions[5].

[5] "The topological relations come out as the strongest discriminators approximately 22 times stronger than all metric parameters combined which confirms the under-

3.1 Methods for the Knowledge Representation Approach

Knowledge is organized in a sample OWL DL Knowledge Base KB, consisting of a TBox T and an ABox A. Partitions of regions are represented in T, partially ordered in a way that each element of a partition is a subset of an element of the next upper level of partitions. Each partition is typed and the concepts for typing are mutually disjoint, so that an individuum can only be of one type. So assume you have the partitions $C(x_i)_{i \in I}$ and $D(y_k)_{k \in K}$ of the same region, their types are C and D. $C(x_i)_{i \in I}$ is understood as more fine-grained than $D(y_k)_{k \in K}$ if each element of $C(x_i)_{i \in I}$ is a subset of an element of $D(y_k)_{k \in K}$. For instance, $District(y_k)_{i \in I}$ is partitioned by elements of $Community(x_i)_{i \in I}$ and both are partitions of a canton. PartOf relations are kept functional, which means that regions are only asserted as part of the next upper region but not as part of the region above the next upper hierarchical step. In the ABox A one finds assertions like *partOf(Dietlikon, District-Buelach)*, stating that the individual Dietlikon, which is of the type community, belongs to Bülach, which is a district of type. Further all individuals are asserted as different from each other. The Region Connection Calculus can be used to represent spatial relations in first order logic. There are different sets available (e.g. RCC-3, RCC-5, RCC-8). For our purpose, we are using RCC-8, which means, we are using a RCC set that differentiates 8 relations. You can see the 8 relations in Figure 1.

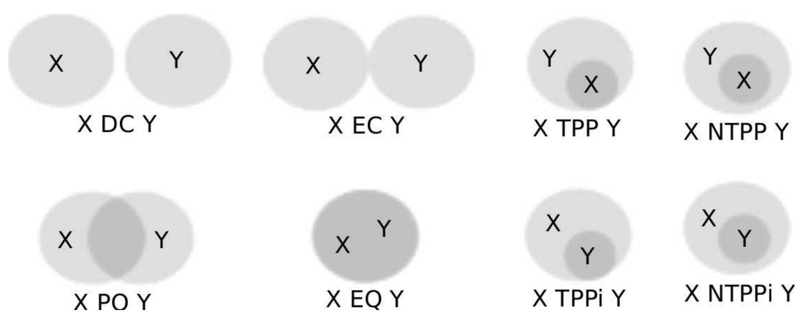

Fig. 1: The Region Connection Calculus with 8 relations

Altogether these relations form a jointly exhaustive and pairwise disjoint set. RCC relations can be interpreted temporal and spatial. Within the spatial interpretion, regions are considered as sets of points. According to that two regions which are connected to each other have at least one point in common. Rules are formulated in a subset of the Semantic Web Rule Language (SWRL)[6]. Our Rule Base is relatively small, in it, existing relations of the Region Connection Calculus (RCC) are used as basis to form Composition Rules. For example there are rules for the additional relation *close to* (cf. [2]).

lying assumption that topology is more critical for the semantics of spatial relations than metric" (Mark and Egenhofer 1994, p. 227 [4]).

From RCC-8, we are using the relations part of P (TPP, NTPP, TPPi, NTPPi), partially overlaps PO and externally connected EC to form the basic relation *close to*. Disjunctions of RCC relations in the bodies of composition rules are represented by auxiliary roles, such as {P, PO} subsuming the roles partOf and partiallyOverlaps. This allows composition rules that are expressed as (non-disjunctive) Horn rules (see equation 1).

3.2 The added Relation *CLOSETO*

To the set of RCC-8 relations we added a composed relation *CLOSETO*. A location x is close to y, stated as *CL(x,y)*. The following equation shows the actual *CLOSETO* rule:

$$\forall x \forall y \forall z [CL_{ap}(y,x) \wedge z\{P, PO\}y \rightarrow CL(z,x)] \quad (1)$$

It is read as region z is close to region x if region y is *a priori* close to x and z is part of or partially overlaps y. This rule makes up the basic building block of this approach. In addition to the basic rule the knowledge representation approach also includes the notion of "a priori"-closeness, which is derived by a second rule (cf. [2]). This second rule enables us to include functional micro regions additionally to the administrative regions into reasoning. These micro regions, consisting of mountane and space planning regions, were introduced to analyze the behavior of commuters. Since these micro regions are also related to traveling their addition seems to be useful for the comparison of the two approaches. For more details please refer to [2].

3.3 Results of the Knowledge Representation Approach

In previous papers it has already been shown that this approach works for the part of Switzerland that is covered by the sample ontology. Also an evaluation has been performed using the search engine "GoForIt"[6], which provides general search and directory search as shown in [2]. For 170 communities two different kinds of queries were performed. Firstly plain queries, such as "communities close to Dietlikon" and afterwards a query which contained a concatenation of all the communities which have been inferred as "close to", such as "Nürensdorf OR Dübendorf OR Rümlang OR Wallisellen OR Kloten OR Wangen-Brüttisellen OR Bassersdorf OR Opfikon" (all communities inferred as close to Dietlikon). Finally, the results showed that recall was significantly higher for the rewritten query.

4 Statistical Approach

We will represent now the second approach which is based on the idea that people speak of places as *close to* if they can reach them quickly. Imagine you

[6] `http://www.goforit.com/`

plan a picnic in a forest nearby and because you have lots of food to take with you you want to go there by car. Then you will speak of a forest you can reach in a reasonable amount of time as *close to*. In terms of metric distance this place could be farer away from your location than another one. But the other place doesn't appear as being *close* to you because it would take longer to go there by car. Sometimes the occurrences of *close to* will differ from one means of transport to another. If you are using a car you can reach things in greater distance easier than while you are walking. On the other hand sometimes you find paths through woods which you can take while you are walking but not if you are driving a car.

4.1 Methods for the Statistical Approach

To gain language data in an appropriate amount the german newspaper text corpora of the Institut für Deutsche Sprache (IDS, Mannheim[7]) were used. Altogether it has a size of 5.3 million tokens. German is used as object language because the application of the approach should start in the German speaking part of Switzerland[8]. The keyword string *in der Nähe von* (i.e. "close to") was looked up in the corpus. Because of the great ambiguity of toponyms and since there are not yet good filters for toponyms available items were annotated manually for the two *close* places. Then all identified place names were looked up in a gazetteer to obtain the coordinates. Geonames[9] and Swissnames[10] were used for this. The additional inclusion of Swissnames aimed at getting rich data of Switzerland so that we could guarantee the comparability to the first approach which is only implemented for a part of Switzerland yet. Then the coordinates of the identified places were fed into a route planner. The routing API of cloudmade[11], which is based on OpenStreetMap[12], served best for this purposes. If the place name was ambigous, the nearest match was chosen. The routes were calculated for trips by car, bike and walking. To get counter examples also hits for "nicht in der Nähe von" (*not close to*), "weitab von" (*further away from*) and other expressions for counterparts of *close to* were annotated. With these there were some difficulties since they are - except the *not close to* - not direct antonyms to *close to* and often they were used as a subjective statement of how far something is away with regard to some topic. Part of the instances, for example *weit entfernt von* ("far away from"), were often used to neutrally express distance in combination with a metric distance measure. An example would be:

[7] `http://www.ids-mannheim.de/`

[8] In other countries other amounts of traveling time maybe felt as *near*. For example, this amount of traveling time for Switzerland may be around 12 minutes. But for larger countries with only few cities this may even be 2 hours. In the future, one could calculate country specific traveling values.

[9] `http://www.geonames.org/`

[10] `http://www.swisstopo.admin.ch/internet/swisstopo/en/home/products/landscape/toponymy.html`

[11] `http://cloudmade.com/`

[12] `http://www.openstreetmap.org/`

10km weit entfernt von ("10km far away from"). And mostly, if places are not near each other, this fact is not explicitly mentioned.

Further there was a technical problem. Sentences like *Frankfurt, weitab von Asien* ("Frankfurt, far away from Asia") occured. One can imagine, that it would not be difficult to calculate a route from Frankfurt to anywhere, but how to manifest the endpoint for that route in a whole continent like Asia? Therefore such routes could not be calculated. Nevertheless, there are some negative examples and calculated trips by car, bike and walking for them as well. Analogous to the "near-matches", when a name was ambigous in geonames, we picked the match with farest distance between the two places.

5 Preliminary Evaluation

The new approach is intended to model *closeness* via reachability with different means of transport (cf. section 4). Reachability is lowered by path obstacles.

5.1 Qualitative Illustration

See Figure 2 for an example where a mountain ridge would force you to make a detour of 3 times the length of the direct path when traveling by car. Because of this detour you would not say that Arosa and Davos, the places in question, are close to each other. The statistical approach can take care of path obstacles like mountains and lakes and missing direct connections between neighboured suburbs and so on. (For the knowledge representation approach these two places would be close to each other when applying the basic rule, because the community of Arosa borders the district of Davos. They would not be close to each other when applying the rule which has been extended for micro regions. Here we have an example where the inclusion of micro regions underlines the reachability.)

5.2 Quantitative Evaluation

For all the 345 pairs of places which occurred in the corpus connected with *in der Nähe von* (English: "close to"), we calculated the route for going by car, going by bike and walking. We collected traveling time and traveling distance. The same holds for the 30 pairs of places which are connected by *nicht in der Nähe von* ("not close to") and the above mentioned (cf. 4.1) synonyms. Then the true distance was also calculated for every pair.

Some data points are quite far away from the rest. A manual check of the 10 sentences in the corpus to the most far away data points showed that there have been mismatches to geonames, because for one of the place names there was only another item in geonames which did not match the place actually meant. Therefore SPSS[13] was used to draw a histogram of traveling time by car only

[13] The SPSS software can be obtained under `http://www-01.ibm.com/software/analytics/spss/`

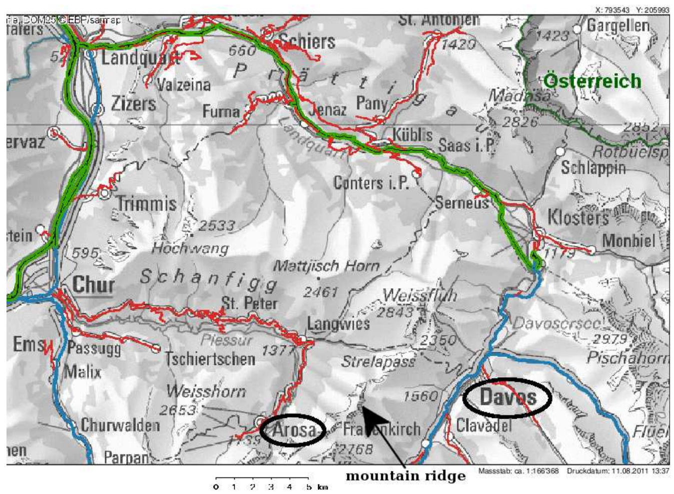

Fig. 2: Arosa and Davosa are departed via a mountain ridge. Since there is no direct street over the mountain, a car has to take all the way around the mountains along Chur, Landquart, Küblis and Klosters. Source of Map: Kantonale Verwaltung Graubünden, GIS-Kompetenzzentrum (http://mapserver1.gr.ch/kantonalesstrassennetz/kantonalesstrassennetz.phtml)

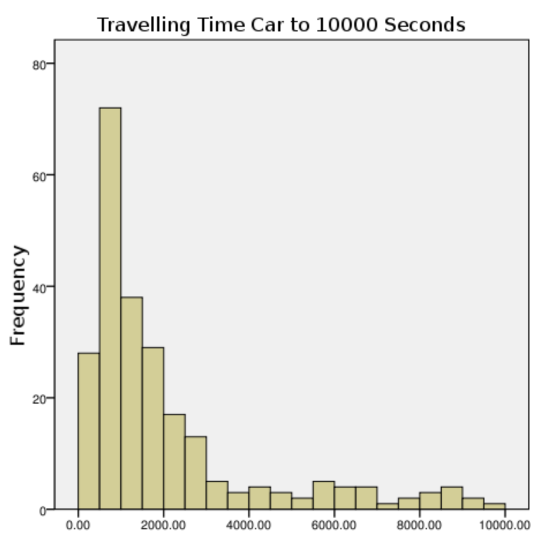

Fig. 3: histogram with 20 equal-distance classes for traveling time up to 10,000 seconds

up to 10,000 seconds. You can see it in Figure 3. It has 20 equal distance classes which have a size of 500 seconds each. This histogram still includes 240 hits. It shows a right-skewed distribution: in the first class we only have 28 occurrences. With 71 occurrences, most points are found in the second class 500 to 1,000 seconds. Afterwards the amount of datapoints declines for the next two classes like normal distributed data. That there are relatively few occurrences in the first class let us conclude that things which are connected to each other are not mentioned as being *close to* each other.

		Number of total Occurrences in Corpus	Knowledge Representation Occurrences (Switzerland only)		Travelling Time by Car in Seconds				
			basic Rule	including Micro regions	Min.	Max.	Median	Average	Standard deviation
All	close to	345	-	-	129	706788	2166	31222	103053
	not close to	30	-	-	9720	605711	34713	166289	215039
Close to <= 10000 Seconds		240	-	-	129	9859	1200	2072	2160
Switzer-Land	close to	33	28	21	265	2382	729	728	485
	not close to	2	2	2	3375	7526	4304	4708	2163

Fig. 4: Overview of Results in Numbers

A table with the results of the comparison can be seen in Figure 4. Like in the histogram, we chose to display the results for the 240 cases up to 10,000 seconds traveling time as well. For example the maximum of time needed for a route to a place would be over 8 days (706788 seconds). As already mentioned this is because of mismatched place names with the gazetteer items. For the same reason average (8.67 hours; 31,222 seconds) and standard deviation (28.63 hours; 103,053 seconds) show high values. The shortest trip to a *closeby* place takes 2,15 minutes (129 seconds). For the 240 cases up to 10,000 seconds of traveling time we have better results. The maximum value is 2,74 hours (9,859 seconds). Average is 34,53 minutes (2,072 seconds). The standard deviation 36 minutes (2,160 seconds) is still high. A reason for this could be that also other means of transport have an influence here. So for example, one feels as *close to* a place where one has a direct flight connection to. This could explain that the histogram declines in waves, the second peak could for example make up the places with direct flight connections. And since the second peak is much smaller, we could conclude that this is because it is much more ususal to go by car than to go by plane. The minimum traveling time of the *not close to* matches is slightly below 10,000 seconds. This may back up that our decision to have a closer look at occurrences below 10,000 seconds. With the *not close to* matches it is quite natural that standard deviation is high, since we also used differnt synonyms for *not close to*. Also the range for things which are felt as *not close to* may be very widespread.

Results for pairs of places in Switzerland are listed again to compare them to the knowledge representation approach which is only applicable to a part of Switzerland right now. Since the knowledge representation approach deals with administrative regions but not with villages we mapped every occurring village to its community for the comparison. For Switzerland we only have 2 negative matches, too few to say much about them. What can be seen by the positive matches is the smaller scale: the range is between 4.42 minutes (265 seconds) and 39.7 minutes (2,382 seconds) and the average for traveling time is only 12.13 minutes (728 seconds). We can conclude from that, that when both places of a *close to* relation are situated in a relatively small country like Switzerland, also the distances between *closeby* places are small.

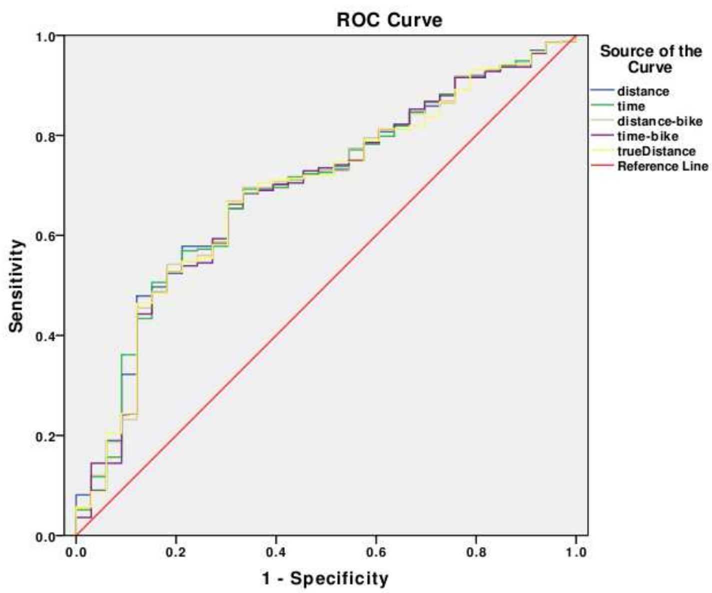

Fig. 5: Sensitivity and Specificity of time (car), distance (car), time-bike, distance-bike and true distance. It can be seen that time (car) and time-bike are slightly better predictors than true distance.

A logistic regression calculation using SPSS with nearness as dependent variable (1 for near pairs; 0 for not-near pairs) showed that distance, traveling distance and time are all correlating to nearness. But still we have too few values for pairs which are *not close to* to do valuable prediction. Also with SPSS sensitivity and specificity were calculated under the assumption that traveling time predicts what is felt as *close to*. Results are shown in the roc-curve in Figure 5. As you can see, the correlation of time was a bit stronger but with the data base available now it shows not significance.

5.3 Comparison with the Knowledge Representation Approach

The comparison with the knowledge representation approach shows, that places that occurred in the corpus as *close to* each other are also found to a great amount by the knowledge representation approach. The best fit was gained with application of the basic rule: 28 of 33 matches were found, which makes 84.85 %. For the micro region extension there have been 21 of 33 matches, which makes up 63.64 %. So the extension with the micro regions is more restrictive and the basic rule found many of the *closeby* places. Maybe in the future one could use additional regions, like the micro regions, to limit results of the knowledge representation approach for a typing structure of regions that is related to the context of searching.

The evaluation is not finished yet, we still have to gain more corpus data for places which are not near. But we already have established the process to get final results for our approach.

6 Discussion

		Knowledge Representation	Statistical
Amount and Type Of Input	+	Satisfied with little data	After establishment: Only route information needed
	-	Needs very special data: Polygons from topologies	Not suitable where no route can be calculated
			Establishment needs much data
Variability of Feature Types For Places	+	Extendable via extension of ontology	Applicable wherever there is a route between two places
	-	Only for entities in ontology (no bakeries, cinemas, villages for now; but extendable)	Data needed as points, e.g. no route for "Frankfurt far away from Asia"
Context For Usage	+	Not as restricted as Statistical Extendedable for further types of regions (e.g. inhabitation density areas)	Best matches for context of travelling with given means of transport
Output	+		Adjustable to vagueness
	-	Binary: either *close to* or *not close to*	

Fig. 6: Overview of Advantages and Disadvantages of both Approaches

The table in Figure 6 shows an overview of the advantages and disadvantages discussed in this section. The knowledge representation approach will be more precise wherever there is only little data about transport connections like streets available. But information about hierarchical structures of regions, information about topology and the type of a region is needed. It is good for modelling all kinds of landscapes (e.g. swamps, mountains) as polygons and reason on them. The statistical approach needs much data for setting up critical times for things which are *closeby*. Once after the approach started working with good predictions it will be applicable wherever you have route information, but not if the route cannot be calculated.

close to-calculations with the knowledge representation approach can be done with all types of places that are specified in the ontology and only these. If one wants to calculate *closeby* cinemas or bakeries, the ontology has to be extended with such types of places and entities which are of these types. An example for this is used in the discussion of the table in Figure 4: some places are villages and only the types of community, district and canton are available. So a mapping of the village to the community to calculate the *close to*-factor for the community in which it was situated was performed.

The statistical approach is applicable whenever there is a known path between places. It works with point data, not polygons. For cities, states and so on there is always used a point which lies within. Depending on where you are in the state/ city the appraoch is more or less accurate. For that reason we have not been able to calculate a route from Frankfurt to Asia (cf. the *Frankfurt weitab von Asien* ("Frankfurt far away from Asia") example from the *not close to* part of the corpus, section 4.1). Nevertheless, under normal circumstances it is not very likely that such a route is needed.

While the Knowledge Representation uses administrative regions, the statistical approach uses the context of traveling. Traveling is important for perception of things *closeby* but also the hierarchical administrative structure of a country has some influence (cf. section 2). The Knowledge Representation method will always lead to clearcut judgements *close to* or *not close to*. But since we are dealing with natural language data which is often quite vague and has many influence factors, in one context a place may be seen as *close to* the reference place whereas in another it is not. The statistical approach can also say something about the "shaded" areas, it can give the degree to which something is near.

7 Conclusion and Outlook

We have shown that it is possible to model human language concepts of spatial relations via description logical expressions using administrative regions as background knowledge or via reachability by different means of transport. They both have advantages and disadvantages. As already mentioned we have to extend the evaluation for the statistical approach. When this is done with satisfying results, we want to embed more possibilities to model humans perception of vague natural language expressions for spatial relations, for example *bei* (English: "next to"), *zwischen* (English: "in between"), etc. We will do this for the two presented approaches and maybe for others which we will develop in the future. Ontologies are providing good background knowledge for such additional models. There are many influence factors for the perception of spatial relation, one could build up a user-friendly system which first evaluates the most important models for the users needs and then computes the best-matching results.

Bibliography

[1] Carbon, C.C., Leder, H.: The wall inside the brain: Overestimation of distances crossing the former iron curtain. In: Psychonomic Bulletin & Review, p. 746750. No. 12 (4) (2005)

[2] Grütter, R., Helming, I., Bernstein, A., Speich, S.: Rewriting queries for web searches that use local expressions. In: Bassiliades, N., Governatori, G., Paschke, A. (eds.) Rule-Based Reasoning, Programming and Applications; Proceedings of 5th International Symposium, RuleML 2011-Europe. pp. 345–359. Springer, Heidelberg (2011)

[3] Maki, R.: Categorization and distance effects with spatial linear orders. Journal of Experimental Psychology: Human Learning and Memory 7 (1), 1532 (1981)

[4] Mark, D., Egenhofer, M.: Modeling spatial relations between lines and regions: Combining formal mathematical models and human subjects testing. Tech. Rep. 94-1, National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA (1994)

[5] Mata, F.: Geographic information retrieval by topological, geographical, and conceptual matching. In: F., I.F., A., R.M., S., L. (eds.) Proceedings of the Second International Conference on GeoSpatial Semantics (GeoS 2007). p. 98113. Springer, Lecture Notes in Computer Science No 4853, Berlin (2007)

[6] Motik, B., Horrocks, I., Rosati, R., Sattler, U.: Can owl and logic programming live together happily ever after? In: Cruz, I.e.a. (ed.) Proceedings of the 5th International Semantic Web Conference (ISWC 2006). LNCS, vol. 4273, pp. 501–514. Springer, Heidelberg (2006)

[7] Randell, D.A., Cui, Z., Cohn, A.G.: A spatial logic based on regions and connections. In: Nebel, B., Rich, C., Swartout, W. (eds.) Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR92). pp. 165–176. Kaufmann, San Mateo, CA Morgan (1992)

[8] Schockaert, S., De Cock, M., Kerre, E.: Location approximation for local search services using natural language hints. International Journal of Geographic Information Science 22 (3), 315–336 (2008)

[9] Yaoh, X., Thill, J.C.: How far is too far? - a statistical approach to context-contingent proximity modeling. Transactions in GIS 9 (2), 157–178 (2005)