

FINAL DRAFT

Cognitive Transactions - A Communication Model

Paul Weiser, Andrew U. Frank

Vienna University of Technology, Austria
Department for Geodesy and Geoinformation
Research Group Geoinformation
Gusshausstr. 27-29, 1040 Vienna
{weiser, frank}@geoinfo.tuwien.ac.at
<http://www.geoinfo.tuwien.ac.at>

Abstract

Whenever a person gets lost and there is no way to access stored spatial information, e.g. in the form of maps, she needs to rely on the knowledge of other humans instead. This situation can be modelled as a communication setting where a person lacking spatial knowledge requests information from a knowledgeable source. The result are cognitive transactions in which information over various levels of detail (LoD) is negotiated. The overall goal is to agree on a shared spatial representation with equal semantics, i.e., a common ground. We present a communication model that accounts for establishing a common ground between two agents. The agents use a modified "wayfinding choreme" language and special signals to negotiate the LoD. Findings of a case study were used to verify and refine our work.

Keywords: Cognitive Transactions, Spatial Negotiation, Level of Detail, Route Descriptions

1 Introduction

Compared to the navigational skills of some animals (c.f. ants, bees, or birds) humans perform rather poorly at this task. As a result, unless specially trained, either formally (e.g., ship navigators) or experientially (e.g., Inuits), we get lost on a quite regular basis [21]. Tools to store and share spatial information, e.g. maps, help us to accumulate expert knowledge with the goal to overcome our inborn cognitive limitations. Also, the manipulation of physical tools reduces the need to carry out computations by the "manipulation of the mind" [15].

There are situations, however, where we cannot access permanently stored information, but have to rely on the knowledge of other humans instead. A prototypical example is that of route descriptions given by another person verbally. They differ from automatically generated or written route descriptions in a number of ways. For example, the information needs to be memorized. Written

instructions can be consulted several times along the way; asking for directions is a single event, bound to the location the communication takes place. Also, verbal instructions frequently combine different levels of detail (LoD), e.g., (1) Walk down the stairs then (2) take train S1; it leaves at 10 a.m from platform A. Note that (1) offers spatial information while (2) offers additional thematic and temporal information.

Allen [3] coined the term cognitive transaction, i.e., the process where one individual (target) seeks information from a knowledgeable source. The LoD at which the source presents information depends, i.a., on the target’s knowledge and the source’s expectation of the target’s knowledge. The target in turn attempts to construct a mental model based on the presented information. Because the source’s and target’s mental models are not necessarily identical, the communication seeks to eliminate the differences. The result is a negotiation over the correct amount of LoD for each part of the route description. The communication ends, if both individuals believe they have established a shared spatial representation with equal semantics, i.e., a common ground (See Section 5).

While there already exists some work on adaptable route descriptions, e.g., in a dialog between humans and machines [29], as well as an analysis of the structure of giving directions ([38], [27]), a comprehensive study of the cognitive processes taking place during a spatial communication setting between two humans is still missing. In this paper, we provide a formal cognitive model of the communication process between an information seeking target and a knowledgeable source. Our model integrates the notion of language advocated by Clark[4], who sees human communication as joint actions between agents, as well as a modified version of "wayfinding choremes" [16]. To achieve this, we introduce special signals used by the agents to indicate granularity changes. Finally, our model makes the differences between the participating agent’s mental models explicit and discusses issues on how to establish a global common ground.

We expect our findings to contribute to the design of better navigation systems, capable of adapting the LoD at which information is presented to a user’s knowledge and possible special needs. A formal understanding of how humans abstract from information and convey it through language allows us to build more intuitive computational models. Language as a tool of study seems particularly promising because it allows us to look into otherwise "hidden highly abstract cognitive constructions" ([7], p. 34).

The remainder of the paper is structured as follows. Section 2 discusses relevant previous work. In Section 3 we briefly review several existing approaches to LoD in route descriptions and show how wayfinding choremes can be expanded to take granularity changes into account. In Section 4, we introduce and discuss various signals an agent can use to indicate a granularity changes. In Section 5, we elaborate on the individual phases during a spatial negotiation setting, as well as the issue of establishing common ground from a global perspective. In section 6, we present the results of our case study used to verify and refine our model. The final section concludes with future research directions.

2 Relevant Previous Work

This section discusses the core concepts relevant to our work. We should note, however, that our approach and choice of terminologies are motivated by the goal to build models that resemble the way the world is perceived and acted upon by ordinary humans. This connects our work to the fields of "Naive Geography" [5] and "Geographic Cognitive Engineering" [28].

2.1 Cognitive Transactions, Wayfinding, and Actions

Navigation can be conceptualized as the combination of wayfinding and locomotion [24]. Lynch [22] defines wayfinding as the "consistent use and organization of sensory cues from the external environment". In terms of Norman [26], sensory cues are "knowledge in the world", i.e., information that can be derived from the world adding to an agent's existing "knowledge in the head". Whether or not wayfinding is successful depends on the two types of knowledge available, i.e., information in the head (mental representation) and information in the world. In case a person has a spatial mental model [36] of the geographic space in question, wayfinding is simply the matching of the actions possible based on the representation in the head with possible actions perceived from the environment (See Figure 1). If the information in the head alone is not sufficient, information from the world can be an additional aid. For example, a street sign could allow me to carry out an action my mental representation would not have allowed. If neither type of information is sufficiently available, a person is considered to be lost.

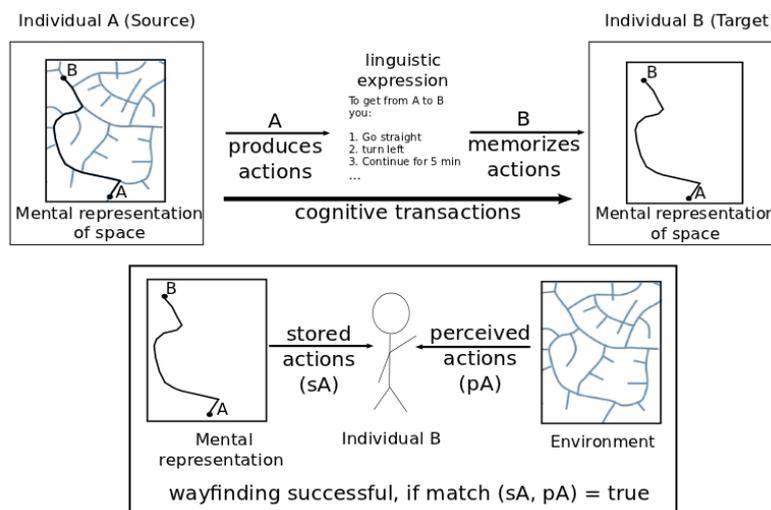


Fig. 1. Cognitive Transactions, Wayfinding, and Actions

In this paper we use the term action to refer to processes performed by an agent. There exists extensive literature concerned with the semantics of process and action. For a recent discussion and comprehensive review of the different approaches see [10]. Galton and Mizoguchi stress the interdependence of objects and processes and advocate to "model reality as it appears to humans engaged in ordinary human activities (p. 3)". From an informal point of view, a route description is a sequence of actions that, if carried out accordingly, allow an agent to navigate from A to B. Generally, we can view the world as having a state at any time. From an agent's perspective, the state of the world can be changed through actions. Actions unfold in space and time and are usually goal-oriented. Therefore, we can conceptualize navigation as a sequence of actions performed by an agent to move from A (source) to B (goal). Furthermore, Allen [3] mentioned the importance of choice points along the route, "affording options with regard to pathways, with intersections being the most typical example".

Klippel [16] provides a formalism of afore mentioned principles in which he states that a route can be seen as a sequence of tuples that consist of a route segment (RS) and a decision point (DP). Here, DPs take the role of Allen's choice points [3], thus they require an agent to decide on a further action, e.g., which direction to take. Klippel [16] also showed that humans have seven different conceptualizations of turning actions at DPs. With these action primitives, or "wayfinding choremes (wc)", we can express most turn-by-turn route descriptions. The following formal definition was adapted from Klippel [16].

$$\begin{aligned} \langle \text{Route} \rangle &::= \langle \text{DecisionPoint} \rangle \langle \text{Segment} \rangle [\langle \text{RoutePart} \rangle] \langle \text{DecisionPoint} \rangle \\ \langle \text{RoutePart} \rangle &::= \langle \text{DecisionPoint} \rangle \langle \text{Segment} \rangle \\ \langle \text{DecisionPoint} \rangle &::= wc_l \mid wc_r \mid wc_s \mid wc_l^{RM} \mid wc_r^{RM} \mid wc_s^{RM} \end{aligned}$$

Note that we use only a subset the original seven actions. We use "turn left", "turn right", and "go straight" which we denote by " wc_l ", " wc_r ", and " wc_s ", respectively. Our definition also includes routemarks (RM), i.e., landmarks along the route [16]. RM are DPs with a special status and we indicate turning actions at RMs by wc_l^{RM} , wc_r^{RM} , and wc_s^{RM} , meaning "turn left at routemark", "turn right at routemark", and "pass routemark", respectively.

2.2 Granularity and Navigation

Our work adds to the growing body of literature on level of detail (LoD) in the context of the communication of route descriptions. Hirtle et al. [12] stressed that the LoD at which information should be presented needs to take the context, more specifically, an agent's activity at hand, into account. Unfortunately, commercial route descriptions only fit for prototypical instances of humans or prototypical activities.

The choice of how much information needs to be communicated can be guided by Grice's [11] conversational maxims. The maxims state that the information communicated should be (1) relevant, (2) not overly redundant but (3) sufficiently detailed for an agent to carry out her task. Sperber and Wilson [32] argue that their concept of relevance can account for all of Grice's maxims. If

the sender tries to be as relevant as possible, the contextual implications that follow reduce the cognitive load for the receiver. It is important to note that two route descriptions, although different in granularity, could both lead to the same outcome. Likewise, two messages of equal granularity may lead to different outcomes, depending on the needs of the user receiving the message. Frank [8] coined the term "pragmatic information content" to account for that fact.

Commercial applications are far from considering afore mentioned requirements. For example, in Figure 2 we see a very detailed route descriptions similar to "Follow x-street for 52 secs, then turn right at intersection y and follow the road for another 23 m". Who would be able to count 52 seconds (clock) or measure 23 meters (odometer) while driving a vehicle?

	18. Turn right onto Südring About 2 mins	go 1.3 km total 180 km
	19. Turn right onto Goethestraße About 52 secs	go 400 m total 180 km
	20. Turn right at Gerhart-Hauptmann-Straße	go 23 m total 180 km
	21. Turn right	go 36 m total 180 km

Fig. 2. Route suggested by Google Maps

An example for a less detailed route description (See Figure 3) is the journey planer offered by the German Railway Company (DB AG). It only lists start- and endpoints of a trip with corresponding times as a function of the transport mode. The actual process of combining those pieces to allow for navigation, i.e., their semantics, is entirely left to the user. Note that we used the same start and end points for both searches (Hamburg Rothenhauschaussee - Rostock Central Station). The route offered by DB AG assumes that we can acquire the missing information "in the world", while Google wants us memorize the route beforehand, so we can access information "in the head" [26]. One could argue that the information offered by DB AG hardly qualifies as a route description in the classical sense. The problem, however, is that people use this information as if it were a route description, mostly because of the lack of better alternatives in this (indoor) context. In general, we can conclude that both descriptions, as well as most automatically generated route descriptions, violate Grice's maxims and do not take the activity at hand into account. Tenbrink and Winter [33], for example, criticize automatically generated descriptions and their "(potentially disruptive) redundancy" as well as their lack of taking prior knowledge into account.

While verbal route descriptions have the potential to be erroneous or too vague to be useful for the receiver, they often make use of an adaptive LoD. For example, Hirtle et al. [13] showed that this is the case for situations that are perceived to be cognitively demanding ("tricky"), e.g. the "absence of appropriate signage or landmarks", as well as complex geometric situations. Tenbrink

Station/Stop	Date	Time / prognosis		Platform	Products
Klgv Rothenhauschaussee, Hamburg	We, 21.11.12	dep	14:28		Bus 228
Bf. Bergedorf, Hamburg	We, 21.11.12	arr	14:40		
walk 5 min.					
Hamburg-Bergedorf	We, 21.11.12	dep	14:45 +0	1	RE 4313
Rostock Hbf	We, 21.11.12	arr	16:51	6	

Fig. 3. Route suggested by Deutsche Bahn AG

and Winter [33] provide empirical evidence that humans adapt the LoD in route descriptions to an addressee’s individual information need. Schwering et al. [31] showed that humans tend to use hierarchically structured route descriptions, using detailed descriptions at decision points and more abstract descriptions for other parts of the route.

2.3 The Structure of Spatial Knowledge

The hierarchical conceptualization of our world (containment relations) becomes evident if one looks at both its spatial and temporal aspects. We often perceive objects as nested within other objects, or as a collection of objects. For example, a tree is an object possibly contained by a collection of trees, called forest. Also, processes can be conceptualized as contained by other processes. For example ”withdrawing money from an ATM” can be seen as part of the process ”applying for a new passport” [1]. Furthermore, empirical evidence suggests that people organize spatial knowledge hierarchically [23] [33], albeit with inherent systematic errors[36]. Timpf et al. [34] proposed a model that accounts for three levels of abstraction of spatial containment relations in the context of planning a trip.

The commonly used metaphor for the storage of spatial knowledge is the cognitive map. As an alternative, Tversky [36] suggested the terms cognitive collages and spatial mental models. Cognitive collages account for the fact that our knowledge is often incomplete and erroneous. As a result not all of its parts can necessarily be integrated. For geographic areas that are well known or simple, Tversky suggests the term spatial mental model. This type of model allows for spatial perspective-taking and inferences; both are crucial for the communication of spatial knowledge. We adapt the notion of spatial mental model for our communication model and discuss its role in Sections 4 and 5.

3 Approaches to LoD in route descriptions

In this section we briefly review common approaches to LoD in route descriptions and specify how the language our agents use to communicate (”wayfinding choremes”) can be modified to take granularity changes into account. For a comprehensive overview on the issues of LoD in route descriptions see [33].

3.1 Segments

Klippel et al. [18] propose a data structure based on "spatial chunking", i.e., the combination of elementary information into higher order elements with the goal to reduce the cognitive load on the wayfinder. This essentially reduces redundancy but also ensures that the pragmatic information content [8] stays equal. Klippel et al. [17] propose rewriting rules (granularity changes) of elementary "wayfinding choremes" to allow for higher-order elements. For example, the following sequence $wc_s + wc_s + wc_r$ ("go straight, go straight, then turn right") can be rewritten as dwc^{3r} ("turn right at the third intersection"). Accordingly, we can group the following sequence $wc_s + wc_s + wc_r^{RM}$ "straight, straight, turn right at landmark" to dwc^{3rRM} ("turn right at routemark"). Note that we deviate from Klippel's original notation to allow for a processing of expressions in both ways (the number of preceding straight segments is saved explicitly in the string). The modified rules can be defined as:

$$(D1) (n*wc_s) + wc_D \longleftrightarrow dwc^{nD}$$

$$(D2) (n*wc_s) + wc_D^{RM} \longleftrightarrow dwc^{nDRM}, \text{ where } n \in \mathbf{N}, D \in \{l, r\}$$

3.2 Decision Points

The chunking method to combine similar route elements into higher-order segments works well with turn-by-turn descriptions. Humans, however, make also use of destination descriptions [35], e.g., "take the train to Leuven, Belgium". Such descriptions are basically discrete 2D containment relations, i.e., "Leuven is contained in Belgium". Destination descriptions emphasize "What is there" rather than "How to get there".

Since "wayfinding choremes" are limited to turn-by-turn instructions we can not account for 2D destination descriptions per se. In our modified language we can, however, indicate granularity changes at decision points with routemarks. Let us assume that we store 3 discrete granularity levels of a routemark in our data model. For example, if we wish to indicate that we refer to a routemark at the most specified level we could say "pass the routemark RM with name N along street S with address A". We codify the most detailed level by " wc_s^{RM} ". Accordingly, "pass the routemark RM with name N at street S" is codified by " wc_s^{+RM} ". Consequently, the most abstract level is indicated by " wc_s^{++RM} ". The following definition shows the granularity change of routemarks (from most specific to most abstract).

$$(D3) "wc_D^{RM}" \longleftrightarrow "wc_D^{+RM}" \longleftrightarrow "wc_D^{++RM}", \text{ where } D \in \{l, r, s\}$$

3.3 Context Specifics

The following subsections briefly discuss context specific approaches to LoD. Their integration into our model, however, is left for further research.

Elaborating on Situations During communication, the amount of knowledge shared between speaker and addressee determines the correct LoD of each utterance. If I come home after work, open the door and say "Hi! I'm home. Where are you?", my girlfriend (if she is there) could answer "I'm in the bathroom" because we share the knowledge that we are both in the same flat. If, on the other hand, I call home from my office and ask her "Where are you?" she will likely be less specific and say "I'm at home" rather than "I'm in the bathroom" because she could be in another bathroom at another flat.

Similarly, if A tells B to "go straight on, until you come to a bank" B could respond "What bank do you mean?" triggering A's possible response "Bank of America". This approach is called elaboration [33] or situational approach to LoD and can be formalized by using partial function application [37]. On a conceptual level this approach is similar to changing the granularity at a routemark but could refer to any entity.

Metonymies In real life we often understand patterns from one domain of experience by projecting it onto another domain of a different kind [20]. We call these mappings metaphors and make use of them very frequently. For example, love is often understood as a journey, e.g., "Our relationship has gotten off the tracks". While metaphors are mappings between two different domains we also make use of cognitive abstraction processes that map within the same domain. For example, the "UK" are often (falsely) referred to by its part "England". The part, i.e., England, stands as a representative concept for the whole UK. A mapping of such kind is called Metonymy.

Humans conceptualize and reason with the help of Metonymies [19]. Naturally, this also applies to the the communication of spatial information. For example, a single mode trip can be divided into phases, each of which is required to successfully complete the trip. According to Lakoff [19], the subprocesses of such a trip can be modeled as follows: In the beginning you are at origin of the trip (Location A). You need access to a vehicle (precondition), then you get into the vehicle (embarkation phase), then you drive (center phase), then at your destination you get off (disembarkation phase), and finally you are at your goal (Location B). Metonymic effects show, for example, when we communicate such spatial descriptions. Imagine you meet somebody at a conference and she asks you "How did you get here?". Consider the following two replies:

1. "First I walked to my car, [...] , I started it up, [...] took a right turn at X, [...] , I parked my car, [...], and here I am"
2. "I drove"

Option 1 includes various descriptions that refer to scripts [30] and are implicitly known to most people as well as a detailed turn-by-turn description of the trip. Violating Grice's conversational maxims and having no relevancy in this context it would bore your friend to death. Option 2 instead, uses the center phase of the model mentioned above to stand metonymically for the entire

trip. Other conceptualizations are possible, for example "I hopped on the train" (embarkation) or "I borrowed my father's car (precondition).

Apart from allowing to abstract from redundant information, Metonymies are a possible source for misinterpretations. If I tell somebody "then you take the 11.15", she will have a hard time unless we have established the shared common ground [4] that we both talk about a specific train to take ("time of motion for an entity involved in the motion"[6]).

4 Signalling between Agents

In this section, we define the notion of signals in the context of communicating route descriptions. Signals are a "method by which one person means something for another"[4]. It is important to mention that signals are not limited to speech alone, but also include gestures, facial expressions, and body language. Thus, the signals presented in this section should not be understood as being bound to a particular modality.

4.1 Principles of Language Use

Our model uses principles developed by Clark [4], who emphasizes the collaborative nature of communication. In his view, the use of language is the exchange of joint actions between a speaker and an addressee. Joint actions are used to achieve a "mutually desired goal", i.e., to solve a particular problem. Clearly, joint actions can only be successful if sufficient information is available to solve the problem. As a consequence, we assume that our source (the agent who provides information) is both willing to share information and has total knowledge of the situation at hand. We can assume this form of knowledge, if we take a spatial mental model [36] for granted.

Joint actions are always taken in respect to each other and try to establish a common ground between the communication partners. Thus, both the speaker's meaning and the addressee's understanding are created from common ground. Clark [4] defines common ground as "the part of information we think we share with somebody else". The attempt to establish a common ground leads us to the notion of agents who negotiate the LoD at which information is communicated.

In this paper, we distinguish between a local and a global common ground. We define local common ground as the information p (e.g. a DP or a sequence of DPs) currently negotiated during a joint action. Since "wayfinding choremes" are actions an agent can carry out at decision points, we can conceptualize local common ground as the actions needed to navigate *one* part of the route, on which both participants agree upon. In contrast, we define global common ground as the set of *all* actions an agent can carry out to navigate the entire route, on which both participants agree upon. Therefore, global common ground can be quantified as the overall differences in the belief systems between the communicating agents (See Section 5). Once agreement on local common ground has been established it becomes global common ground.

4.2 The Source's and Target's signals

In Section 2 and 3 we established a simple language, i.e., an extension to "wayfinding choremes", our agents can use to communicate route descriptions over multiple LoD. We can now define the signals that allow both source (S) and target (T) to negotiate the LoD. The signals were extracted from the interviews we conducted during this research (See Section 6). The open-ended records were coded based on the approach described in Montello [25].

Present A basic signal for S to present information (a particular DP or a sequence of DPs) at a given LoD (the LoD need not to be constant and can vary).

Accept A basic signal for T to accept the piece of information (a particular DP or a sequence of DPs) at a given LoD presented by S.

Probe A basic signal for S to probe information from T. Probing can refer to a particular DP ("Do you know xy-routemark?") or to a sequence of DPs ("Do you know how to get from DP₁ to DP₂?")

Reject A basic signal for both S and T. If used by T as a response to S probing information it means that T does not know the DP or a sequence of DPs suggested by S. If used by S as a response to T requesting an LoD change it means that S cannot provide the requested information.

Secure A basic signal for T to make sure information from S was understood correctly. T's secure signal can refer to a DP or a sequence of DPs.

RequestLoDChange A basic signal for T to explicitly request a change of LOD (of a particular DP or a sequence of DPs).

LoDChange A basic signal for S to adjust the LoD of a DP or a sequence of DPs. This can happen explicitly, e.g., as a response to T's request for a change of LoD, but also implicitly, e.g., if T wants to secure information and S reacts to it (See signal combinations in next subsection).

OfferChoice A basic signal, S can use to offer T a choice on alternative paths T could take.

4.3 Signal Combinations

In this section we elaborate on some typical signal combination we observed during the interviews, and apply them to example utterances made by source (S) and target (T) in our modified "wayfinding choreme" language.

S:Present and T:Accept This is the simplest signal combination. S presents information and T accepts it. The information is added to the global common ground of both participants and thus minimizes differences between both agent's belief systems (See Section 5). More formally, S knows that T knows that p and T knows that p , where p is the currently negotiated part of the route.

Note that all other signal combinations need an implicit accept signal to indicate the end of the negotiation sequence (p can then be added to the global common ground). Present does not imply that the information presented is at a constant LoD (See EX 1). In fact, an agent may use a varying degree of LoD in the same sequence.

(EX 1) S: Present $[dwc^{2l}, wc_s, wc_s] \longrightarrow T : Accept$

S:Present and T:RequestLoDChange We can distinguish LoD changes from specific sequences to abstract chunks (EX 2), LoD changes from abstract chunks to specific sequences (EX 3), and LoD adjustments at a decision point with a routemark present (EX 4). Other combinations are possible but not listed here explicitly.

(EX 2) S: Present $[wc_s, wc_s, wc_r] \longrightarrow T : RequestLoDChange \longrightarrow S : LoDChange + Present[dwc^{3r}]$

(EX 3) S: Present $[dwc^{2l}] \longrightarrow T : RequestLoDChange \longrightarrow S : LoDChange + Present[wc_s, wc_l]$

(EX 4) S: Present $[wc_r^{RM++}] \longrightarrow T : RequestLoDChange \longrightarrow S : LoDChange + Present[wc_r^{RM+}]$

T: Secure and S: (Accept+LoDChange) / Reject If T tries to secure an information presented by S she usually wants to make sure whether the information was understood correctly and can be added to the global common ground. S can reject the secure signal (EX 6), i.e., the utterance was not correctly understood. Alternatively, S can accept the securing attempts by T. S can also accept the secure signal and repeat the same information at a higher LoD to remove potential ambiguities (EX 5).

(EX 5) T: Secure $[dwc^{2l}] \longrightarrow S : Accept + LoDChange + Present[wc_s, wc_l]$

(EX 6) T: Secure $[wc_s, wc_s] \longrightarrow S : Reject + Present[wc_s, wc_l]$

S:Probe and T:Accept / Reject In case S probes information p and T accepts the probing, S has successfully determined that p is part of their common ground. For example, if S establishes a sequence of DPs to be common ground for both S and T, the negotiation could skip this part of the route and continue at the first DP after the probed sequence. In case a probing is rejected S can not assume that p is part of the common ground of both S and T (EX 7).

(EX 7) S: probe $[wc_s^{RM}] \longrightarrow T : reject[wc_s^{RM}]$

Other combinations are possible, e.g., S:Probe and T:requestLoDChange, or T: RequestLoDChange and S:Reject, but not discussed here in detail.

5 The negotiation phases

In Section 3, we established a language agents can use to communicate route descriptions. In Section 4, we introduced various signals that can be used to adjust the LoD of such descriptions. In this Section, we discuss possible phases during the negotiation process and their effect on the spatial mental representations, i.e., the global common ground.

5.1 Motivating Example

Imagine Alice, who is a tourist, visiting Vienna for the first time. After a day of sightseeing she suddenly realizes that she is lost. She is under time pressure because the last train with destination to her home town leaves in one hour. She stops Bob on the street, who turns out to be a local, and asks for the fastest way to the central station.

5.2 Initial Situation

In the following we use the abbreviations N and M to refer to Alice’s and Bob’s spatial mental model, respectively. Alice’s initial situation is as follows: She can neither utilize her ”knowledge in the head” (N(Alice)) nor can she acquire the necessary information available in the world to carry out the actions (actions(N)) necessary to find her way to the central station (DP Goal).

$$actions(N) = DP, DP \neq Goal \quad (1)$$

Bob’s situation looks different: He has the necessary knowledge in the head (M(Bob)) and can carry out actions (actions(M)) to get to the central station (goal state).

$$actions(M) = DP = Goal \quad (2)$$

The initial situation is characterized by the fact the Bob and Alice share none or little common ground. To quantify the notion of common ground we conceptualize spatial mental models as the set of all actions that be carried out on the structure of the environment. In terms of both Bob’s and Alice’s mental models, this can be described as follows:

$$|\Delta(M, N)| > 0 \quad (3)$$

5.3 Theory of the Mind

How does Bob communicate the correct amount of information to Alice? In order to follow the conversational maxims, i.e., to maximize the relevancy of his message, Bob needs to put himself into the shoes of Alice. Knowing (rather the belief to know) what the other person knows can help to choose an initial LoD.

The ability to impute mental states to oneself and others has been termed "theory of the mind". This includes the notion that my own mental states can be different to somebody else's. Grice (1989) makes use of this concept in his example of how pragmatic inferences during a conversation (implicature) are made:

"He has said that p; [...]; he could not be doing this unless he thought that q; he knows (and knows that I know that he knows) that I can see that the supposition that he thinks that q is required; he has done nothing to stop me thinking that q; he intends me to think, or is at least willing to allow me to think, that q; and so he has implicated that q"(p. 31, emphasis by the authors)

Empirical evidence presented by Fussell and Krauss [9] suggests that one takes other people's knowledge into account when communicating a message. As a result, people design the communication of their knowledge depending on the audience ("audience design hypothesis"). More recently in the context of route descriptions, Hoelscher et al. [14] concluded that written route descriptions, as well as the actual traveled route are different when presented to somebody else, as compared to one's own conceptualization of the same route.

While the actual process of "extracting" somebody's knowledge is poorly understood, we assume that the image of the mental states of another person is influenced by many factors. Some include (1) The language spoken: Can the person read signs and therefore extract knowledge from the world?, or, (2) Special needs: The elderly, disabled persons, or children may require additional information based on accessibility and/or safety. See Section 6 for responses of what the participants in our case study thought they had designed explicitly for the target.

We can conclude that Bob's expectation of Alice's knowledge results in an image of Alice's mental model in Bob's mind (N_M , see Figure 4). Bob has to deal with two different belief systems, his own, and one he thinks resembles Alice's. Because the theory of the mind can only account for an approximation of Alice's actual model we express this fact in Equation 4.

$$|\Delta(M, N_M)| \neq |\Delta(M, N)| \quad (4)$$

5.4 Negotiation

This phase makes use of the signals mentioned in Section 4.2. Because we assume Alice's knowledge and Bob's image of Alice's knowledge (global common ground) not to be equal, the negotiation seeks to eliminate or at least minimize the differences. Bob needs to recall his knowledge and describe it to Alice verbally at a given LoD (signal present). The LoD chosen can vary for each part of

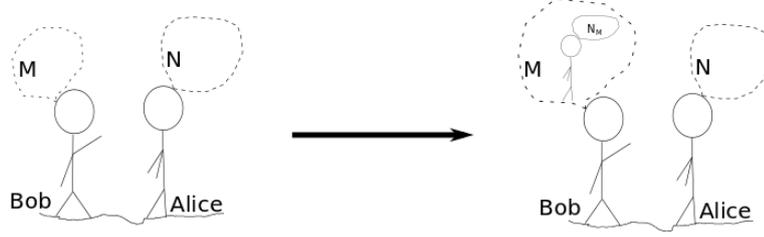


Fig. 4. Applying the Theory of the Mind

the description and depends on Bob's own mental representation of geographic space, his expectation of Alice's knowledge, and his own perceived difficulty of a certain section of the route resulting in an emphasized treatment of that particular "tricky part" [13].

Also, Bob may repeat certain sections of the route description as well as increase or decrease its LoD (present signal), depending on Alice's reactions to his elaborations (accept, or, secure signals). Alice needs to process the description and memorize it (build her mental model) and signal Bob whether the information he presents is either too little or too much detailed (requestLoDChange signal). In any case, the negotiation of a route part with the help a joint action has the goal to agree on a local common ground.

The effect of probing on the common ground During probing (signal probe), Bob attempts to update his image of Alice's model through inquiring on knowledge on a particular DP or a sequence of DPs, similar to "Do you know how to get to Karlsplatz?". The goal of probing is to minimize the differences between Bob's mental model and his image of Alice's model, i.e., to increase the global common ground. Figure 5 illustrates the probing of knowledge by Bob (f = probe signal) and Alice's answer ($f(k)$ = accept/reject signal) as well as its effect on Bob's mental image of Alice's model (N'_M).

$$|\Delta(M, N'_M)| < |\Delta(M, N_M)| \quad (5)$$

The effect of knowledge presentation on the common ground Instead of probing, Bob may present knowledge (signal present) to Alice at a given LoD for each section of the route description (See Figure 6). The effect of Bob presenting his knowledge is an updated mental model for Alice (N'). Knowledge presentation reduces the differences (See Equation 6) that existed between Bob's and Alice's initial representations (M, N) and the updated representations (M, N'). This results in adding the route part both agree upon (local common ground) to be added to the global common ground.

$$|\Delta(M, N')| < |\Delta(M, N)| \quad (6)$$

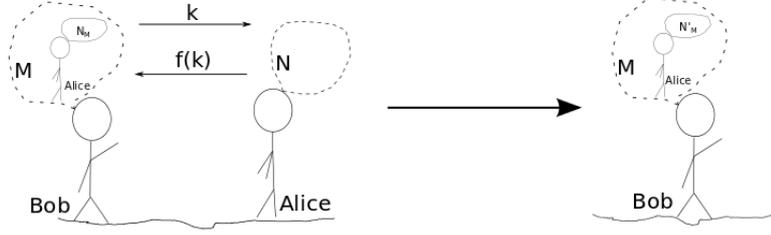


Fig. 5. Probing Information

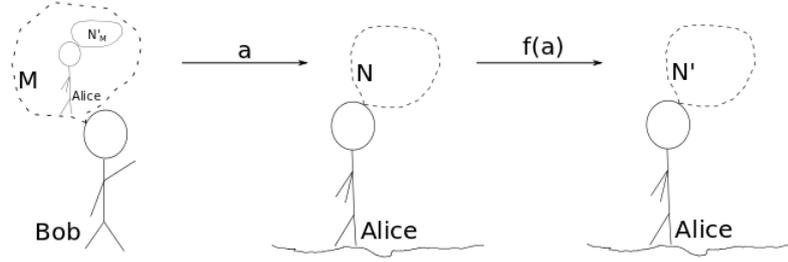


Fig. 6. Bob's knowledge presentation and its effect on Alice's model

5.5 Final Situation

The communication ends once Bob has finished his description of the route, and both participants think Alice has the relevant knowledge to successfully complete her wayfinding task. In other words, this is the case if the predicted actions of Alice $p(N')$ match the predicted actions of Bob's image of Alice's model ($p(N_M)$).

$$p(N_M) = p(N') = DP = Goal \quad (7)$$

Note that we cannot say that Alice's actions lead to the goal state (actions(N') = goalState) because we cannot be certain that Alice will find the way previously described to her. Indeed, the model rather describes both Bob's and Alice's beliefs, i.e., what they predict will happen. Both participants believe to have established an equal global common ground. In terms of both mental models the (optimal) final state can be described as follows:

$$|\Delta(M, N')| \simeq 0 \quad (8)$$

6 Case Study

In this section, we discuss the results of our case study, analyzing the interactive communication of route descriptions between participants and experimenter.

6.1 Design

We asked 10 participants (5 female, 5 male) with a mean age of 31.2 to take part in our case study. The participants were asked to describe a route they travel frequently and know well (e.g. from work to home). This particular set-up ensured that participants were likely to have a spatial mental model which allows for perspective taking, inferences about spatial locations, and "total" knowledge of the situation [36]. The route description was meant to be interactive, i.e., the experimenter took the part of the target lacking the knowledge.

6.2 Findings

The use of different approaches to LoD Most participants provided detailed turn-by-turn instructions throughout the entire route description. 5 of 10 participants made use of destination descriptions, mostly once the description was about to reach the end of the route. Spatial Chunking was used by all but one participants. Participant I (See Table 1) used spatial chunking in 5 of the 17 joint actions. Elaboration on specific DPs happened as a response to requests on LoD Change or as a response to securing signals (See next subsection for a more detailed discussion).

Metonymies were only used by four participants, mostly to indicate line numbers of buses, or subways, or to conceptualize entire route parts with the help of one particular DP, as the following example shows. The route description starts by mentioning to use a subway line (present signal). A means to say "my route description starts at Karlsplatz" using the subway to conceptualize Karlplatz. We should note, however, that the route description should start at Gusshausstrasse (the place the experiment took place), which is about 5 minutes on foot from Karlsplatz. B's attempt to verify this fact (secure signal) triggers an increase of level of detail (loDChange + present signals) in A's repetition of the first section of the route.

A: I only need one public means of transport, (line) U4 (*Ich brauch nur ein öffentliches Verkehrsmittel, das ist die U4*)

B: So, you walk from here to Karlsplatz? (*Also, du gehst hier zum Karlsplatz?*)

A: I walk to Karlsplatz, through Resselpark, then I get on U4 (*Ich geh zum Karlsplatz, durch den Resselpark, dann steig ich in die U4*)

Another observation we made is the explicit referral to knowledge in the world. Instead of giving detailed descriptions for three potentially difficult decision points, one participant referred to knowledge that can be extracted from the world. He recommended to ask the bus driver for directions, as well as to check the time tables of trains because he was not entirely sure what train one should take. Another participant explicitly referred to signs visible in the train station, i.e., "then follow the signs...".

The use of signals and their combinations The following table (Table 1) shows the results of our analysis on the total number of joints actions per participant (the number of signal combinations until local common ground was achieved) and the most prominent signal combinations. Signals could also be seen from a more detailed point of view, e.g., we differentiate between "secure DP" and "secure sequence of DP". The table, however, does not show this differentiation and groups all sub-signals together.

We can see that an explicit request for increase of LOD happened relatively seldom, with exception of participant C, E, and, H, who used it in every fourth joint action. On a more detailed level, however, we noticed that the target's attempts to secure a particular DP or a sequence of DPs almost always resulted in attempts by the source to present the information again, but now at a higher LoD. Apparently, the target's attempts to secure information is perceived by the source as an implicit request for LoD change. We also noticed that a securing signal by T is most of the time accepted by S, indicating that T's understanding of S's meaning was correct. Choices on alternative paths were offered by 5 participants. Only 3 participants used some form of probing.

If we look at how often the present-accept signal combination (T was sure to have received enough information) was used in comparison to the other signals (T was not sure and had to request more information) we could deduce the following: The routes offered by A and I were perceived as the most difficult ones (both about 4x more "insecurity" than present-accept signals), whereas routes described by D and, F were perceived as the easier ones (both about 1.5x more present-accept than "insecurity" signals).

Table 1. Total no. of joint actions and signal combinations per participant

Participant	Joint Actions	Present-Accept	Secure-Accept	Secure-Reject	Request LoDChange	Probing	Offer Choice
A	12	2	6	2	2	0	0
B	16	7	6	2	0	1	0
C	18	6	4	3	4	0	1
D	25	15	5	2	2	0	1
E	15	5	5	0	4	0	1
F	32	17	9	1	1	3	1
G	18	5	7	1	3	1	1
H	19	6	8	0	5	0	0
I	17	4	8	3	2	0	0
J	35	20	8	3	4	0	0

Theory of the Mind We expected the participants to adjust the message to the needs of the receiver as suggested in Section 5. Because this process happens subconsciously, we asked the participants after they had delivered their

description what they assumed the other person could have known and whether this had influenced their description.

As expected, participants referred to the use of routemarks (easily recognizable) and said they tried to keep the description very general. Some mentioned they would have adjusted the description accordingly, if they had communicated the route to somebody who was not familiar with Vienna at all (e.g. foreigners). Parts of the routes participants perceived as difficult themselves were given special treatment leading to more elaborated descriptions. Although participants thought they used very general descriptions, they often used semantics that were not entirely clear to the experimenter (e.g., what does "towards the city limits" mean? Possible interpretations: Right, Left, North, South, etc.). This suggests a gap between one's own perception of space and the audience design hypothesis mentioned in Section 5, a fact already recognized by Hoelscher et al. [14].

7 Conclusions and Future Research

We presented a model that accounts for cognitive transactions in a communication setting between a knowledgeable source and an individual lacking spatial knowledge. Agents use different LoD for the presentation of information and this fact is reflected in the language they use. We modified "wayfinding choremes" to account for granularity changes and developed signals the agents can use to indicate such adjustments. The overall goal of negotiating the LoD is the agreement on a common ground. We discussed the differences between local and global common ground, elaborated on the various phases during a spatial negotiation setting, and made their effect on the common ground of agents explicit.

We argue that our investigation can be a fruitful approach to design better navigation systems, capable of taking the user's knowledge and special needs into account. Commercial applications have, so far, failed to achieve this goal. Some directions for future work include:

Prototype Application A short-term goal of our research is the design, of a prototype application capable of simulating the interactive communication process between a computer and a human agent. The model presented in this work serves as the conceptual foundation for its development. To improve the model presented here, it will be necessary to repeat the experiment described in Section 6 on a larger scale. We expect to identify more signals and a more fine-grained view on the various signal combinations.

Language Extension Klippel [16] noted that wayfinding is a goal-oriented task. We speculate on a simple language that is capable to communicate spatio-temporal tasks in general. This could contribute to the design of better (spatial) personal information management tools [2], e.g., calendars with a sense of space.

References

1. Abdalla, A., Frank, A. U.: Combining trip and task planning: How to get from A to passport. In: Xiao, N., Kwan, M.P., Goodchild, M. F., and Shekhar, S. (eds.), GIScience 2012, LNCS, vol. 7478, pp. 1–14. Springer, Heidelberg (2012)
2. Abdalla, A., Weiser, P., Frank, A.U.: Design Principles for Spatio-Temporally Enabled PIM Tools: A Qualitative Analysis of Trip Planning. In: Vandenbroucke, D., Bucher, B., Crompvoets, J. (eds.) AGILE 2013, LNGC, pp. 323-336, Springer (2013)
3. Allen, G.: From knowledge to words to wayfinding. In: Hirtle, S. and Frank, A. (eds.), Spatial Information Theory: A Theoretical Basis for GIS, LNCS, vol. 1329, pp. 363–372. Springer, Heidelberg (1997)
4. Clark, H.H.: Using Language, Cambridge University Press (1996)
5. Egenhofer, M.J., Mark, D.M.: Naive geography. In: Frank, A. U., Kuhn, W. (eds.) Spatial Information Theory: A Theoretical Basis for GIS, LNCS, vol. 988, pp. 1–15. Springer, Heidelberg (1995)
6. Evans, V., Green, M.: Cognitive Linguistics - An Introduction. Edinburgh University Press (2006)
7. Fauconnier, G.: Mappings in Thought and Language. Cambridge University Press (1997)
8. Frank, A. U.: Pragmatic Information Content: How to measure the information in a route description. In Perspectives on Geographic Information Science, pp. 47–68, Taylor & Francis (2003)
9. Fussell, S.R., Krauss, R.M.: Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. In: Journal of Personality and Social Psychology. 62 (3), pp. 378–391 (1992)
10. Galton, A., Mizoguchi, R.: The water falls but the waterfall does not fall: New perspectives on objects, processes and events. Applied Ontology, 4(2), pp. 71–107 (2009)
11. Grice, P.: Studies in the Way of Words. Harvard University Press (1989)
12. Hirtle, S. C., Timpf, S., Tenbrink, T.: The effect of activity on relevance and granularity for navigation. In: Proceedings of the 10th international conference on Spatial information theory, COSIT 11, LNCS, vol. 6899, pp. 73–89, Springer, Heidelberg (2011)
13. Hirtle, S.C., Richter, K-F., Srinivas, S., Firth, R.: This is the tricky part: When directions become difficult. In: Journal of Spatial Information Science, 1, pp. 53–73 (2010)
14. Hoelscher, C., Tenbrink, T., Wiener, J.M.: Would you follow your own route description? Cognitive strategies in urban route planning. In: Cognition, 121 (2), pp. 228-247 (2011)
15. Hutchins, E.: Cognition in the Wild. The MIT Press (1995)
16. Klippel, A.: Wayfinding choremes: Conceptualizing wayfinding and route direction elements, PhD Thesis, SFB/TR 8 Spatial Cognition (2003)
17. Klippel, A., Tappe, H., Kulik, L., Lee, P.: Wayfinding Choremes - A language for modeling conceptual route knowledge. In: Journal of Visual Languages and Computing, 16 (4), pp. 311–329 (2005)
18. Klippel, A., Hansen, S., Richter, K.-F., and Winter, S. (2009). Urban granularities. A data structure for cognitively ergonomic route directions. In: Geo Informatica, 13, pp 223–247, (2009)
19. Lakoff, G.: Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. University of Chicago Press (1987)

20. Lakoff, G. and Johnson, M.: *Metaphors We Live By*. University of Chicago Press (1980)
21. Levinson, S.: *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge University Press (2003)
22. Lynch, K.: *The Image of the City*. The MIT Press (1960)
23. McNamara, T.P., Hardy, J.K., Hirtle, S.C.: Subjective Hierarchies in Spatial Memory. In: *Journal of Environmental Psychology: Learning, Memory, and Cognition*, 15 (2), pp. 211–227 (1989)
24. Montello, D. R.: *The Cambridge Handbook of Visuospatial Thinking*. Cambridge University Press. (2005)
25. Montello, D.,R., Sutton P.,C.: *An Introduction to Scientific Research Methods in Geography*, Sage Publications (2006)
26. Norman, D. A.: *The Design of Everyday Things*. Doubleday Books (1988)
27. Psathas, G.: Direction-Giving in Interaction. In: *Reseaux* (8), pp. 183-198 (1990)
28. Raubal, M.: Cognitive engineering for geographic information science. In: *Geography Compass*, 3(3), pp. 1087–1104 (2009)
29. Richter, K.-F., Tomko, M., Winter, S.: A dialog-driven process of generating route directions. In: *Computers, Environment and Urban Systems* (32), pp. 233-245 (2008)
30. Schank, R. C., Abelson, R. P.: *Scripts, Plans, Goals, and Understanding* (1977)
31. Schwering, A., Li, R., Anacta J.,A.,A.: Orientation Information in Different Forms of Route Instructions. In: *Short Paper Proceedings of the 16th AGILE Conference on Geographic Information Science*, Leuven, Belgium (2013)
32. Sperber, D. and Wilson, D.: *Relevance : Communication and Cognition*. Basil Blackwell Oxford (1986)
33. Tenbrink T., Winter S.: Variable Granularity in Route Directions. In: *Spatial Cognition and Computation*, 9, pp. 64–93 (2009)
34. Timpf, S., Volta, G. S., Pollock, D. W., Egenhofer, M. J.: A conceptual model of wayfinding using multiple levels of abstraction. In *Proceedings of the International Conference GIS - From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning on Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pp. 348–367, Springer, Heidelberg (1992)
35. Tomko, M., Winter S.: Pragmatic construction of destination descriptions for urban environments. In: *Spatial Cognition and Computation*, 9 (1), pp. 1–29 (2009)
36. Tversky, B.: Cognitive Maps, Cognitive Collages, and Spatial Mental Models. In : *Spatial Information Theory: A Theoretical Basis for GIS*, LNCS, vol. 715, Springer, Heidelberg (1993)
37. Weiser, P., Frank, Andrew U.: Process Composition And Process Reasoning Over Multiple Levels Of Detail. In: *Online Proceedings of the 7th International Conference GIScience*, Columbus, Ohio (2012)
38. Wunderlich, D., Reinelt, R.: How to get there from here. In: Jarvella, R.J., Klein, W. (eds.). *Speech, Place, and Action*, John Wiley & Sons (1982)