Why Context Matters in Giving Driving Instructions —The Necessity for Representing Decision Points

Position paper for workshop "Spatial Behavior and Linguistic Representation"

Andrew U. Frank

TU Wien, Geoinformation,
Gusshausstrasse 27-29/E127.1,
A-1040 Vienna, Austria,
frank@geoinfo.tuwien.ac.at

Abstract

A critical review of past research can lead to identify new, fruitful research directions: here I consider the research on driving instruction by linguists and, especially, by Geoinformation scientists, which has concentrated on static spatial situations. The insight obtained from qualitative spatial reasoning is limited because it did not include actions, change, and motion.

The important use case for dynamic spatial behavior is navigation: how are instructions communicated to an actor, e.g., from the car navigation system to the driver and how are they acted upon? An analysis of the communication situation must be dynamic; while moving the driver identifies "decision situations" at points where instructions are expected. The participants in way finding communication must have the "decision situations" as shared context. Maps represent static spatial situations (Kuhn, 2010) and miss the dynamic situations a navigator encounters. Research on recognition of decision situations is recommended to identify decision situatins for the navigation system, which are the context which the navigator assumes to be shared.

1 Representation of Static Spatial Situation

Classic research in qualitative spatial reasoning studied natural language methods to communicate spatial situation Levinson (1996); it focused on sentences like "the ball is in front of the tree" and translate them to spatial qualitative representation like "front (Ball, Tree)". Different perspectives and how to convert between different perspectives were discussed (Klatzky, 1998)(Frank, 1998). Formalists derived methods to obtain logical conclusions from multiple relations (Freksa, 1991)(Frank, 1992), but not all results were meaningful. The problems posed by imprecise descriptions, especially for distances, limited the expressiveness of results of relation composition (Bird and de Moor, 1997)(Schröder, 1890), a point already made by Egenhofer and Mark in 1995. Talmy had stressed the importance of considering dynamic spatial situation (1990) and introduced "fictive motion" as a device where language casts a static situation to a dynamic one in order to focus on relevant aspects; e.g. "the road runs along the river".

Formalist representation of movement translates movement to vectors, but seems not to capture the essence of moving, e.g., from home to work (Güting and Schneider, 2005). I am not aware of much qualitative dynamic spatial reasoning research (except for (Hägerstrand, 1967)Medak (1999b)Medak (1999a)(Hornsby and Egenhofer, 1997)); perhaps the results from qualitative differential equation are applicable Kuipers (1994)(Kuipers, 1994).

2 Use Case

I believe that use cases are crucial for successful research. Wayfinding is the prototypical example of spatial behavior and computerized devices to produce navigation instructions have become one

Fig. 1: a) left turn, b) keep left

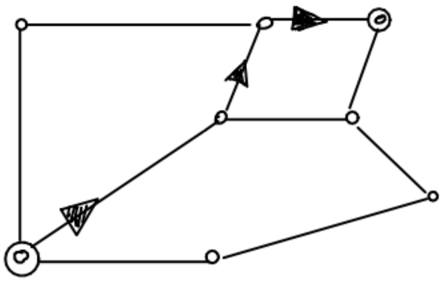

Fig. 2: Shortest path from start to target

of the great commercial successes of the geoinformation industry. Analyzing the current state of the art and its limitations can lead to new research directions to deal effectively with the impediments of curent technology. I will assume—to focus on the issues of interest at the workshops and not to get lost in technology discussions—that position is correctly determined and the stored maps are up-to-date; my interest is in communication and imperfect results are caused by communication failures.

Past research focused what is the most natural way to instruct drivers, i.e. what is the best vocabulary and terminology.

From my own experiences I conclude that naturalness of instructions is not as important as is consistency; I am willing and able to adapt to whatever coding is used in the verbal instructions I am given to encode spatial behavior, e.g., "turn right" vs. "keep left" (Fig. 1).

3 Wayfinding Theory

The theory behind car navigation is very simple: the street network is represented as a located graph. The positioning device maps the current position to a point on the graph. An optimal path to the target is determined by a shortest path algorithm (Dijkstra, 1959). The optional path determined is represented as a list of edges to traverse (fig. 2). At each node in the graph the next edge to select is given; thus at the nodes, where the driver must make a decision and select one of the edges leaving the node, effective advice is expected. – The navigation is successful when the driver translates the instruction in appropriate behavior that guides the vehicle along the optimal path.

4 Pragmatics of Wayfinding

The pragmatics of the information the driver receives are the actions, into which the driver translates the information received. He expects and requires instructions before he chooses between different possible actions. I will call this a *decision situation* d indexed by locations s and a set of possible actions of which a_n one must be selected:

$d_s = \{a_1, a_2, ..., a_n\}$

The instruction given at the location s is i_s and should be one of the a_n at location s, to indicate the action to preform in order to follow the optional path.

5 Failure Modes

What can go wrong? The driver must relate instruction i_s to exactly one of the actions in d_s, say a_k, and then execute the action. The driver, given his perception of the decision situation

Fig. 3: a) the action from the instruction cannot be executed,
b) the action is not matched for difference in classification

Fig. 4: a) the decision point to which the instruction applies is not perceived,
b) the instruction is applied to early

$d_s = \{a_1, a_2, .., a_n\}$ and instruction i_s may succeed in matching one a_i with the i_s or the navigation fails at this node; different cases of failures can be separated:

1. The driver cannot match any action a_i to the instruction i_k (Fig. 3) , because (1a) the action indicated by the instruction i_k is not perceived as a possible action by the driver (i.e. not included in $d_s = \{a_1, a_2, .., a_n\}$; or because (1b) the instruction i_k and actiona_i are not found to match by the driver. The later case (1b) is a proper failure in understanding instruction: The driver cannot use the information in the instruction to select an action for differences in the expression expected and use. The first case (1a), however, is one of oversight: the road to follow is simply not seen.

2. The instructioni_k does not apply to decision situation d_l (Fig.4). The confusion can be caused by the driver perceiving a decision situation at l and expecting instructions for it, which the system does not contain (failure mode 2a - commission) or by a driver not considering a situation as a proper decision situation and not expecting instructions (2b - omission); as a result the instruction received for k will be applied to the next decision situation $k + 1$. In both cases, the instruction is applied to the wrong situation; (2a - commission) the instruction is applied to an situation earlier whereas in (2b - omission) to a situation later.

   In both sub-cases of failure mode 2, the information i_k for location k is understood by the driver to instruct him to select an action in decision situation d_1 at location l. Given that the vocabulary of instruction is very small, the instruction likely appears applicable and is followed, only to be discovered seconds later as wrong (my car navigation device then instructs me to turn around!).

6 Issue related to Failure Mode 1: Matching Instructions and Actions

The driver must match the instruction against his encoding of the visually perceived environment; the driver's encoding is in terms of possible actions which appear possible to him. The encoding used in the instructions follow from the classification used during data collection for the navigable map (White, 1991). Differences between the classification leads to failing matching the instruction to a possible action. Failures often relate to whether a drivable path is considered a street and instructions relate to it or is just an entrance, for example for a shopping center, or a driveway to a building and thus not referred to by driving instructions.

Current technology encodes spatial situations as located graph and deduces from this the required changes in direction (e.g., "turn sharp right"). Difficulties emerge when the driver must select between several options that are taking similar directions; common language does not have devices to quickly identify one of several exits from, e.g. a roundabout Adding symbolic 3D representations (i.e., a reconstructions from the map) of the situation is of limited utility as they do not optically match reality that is visible and may confuse in their lack of matching the visual environment truthfully.

7 Issue related to Failure Mode 2: "Missing the Turn"

The second failure mode is misunderstanding where an instruction applies. An instruction for one location is applied to another location. The navigation system identifies decision situations by geometric location, the driver by visual perception. Differences occur :

- Failure mode 2a: when user perceives a decision situation when the navigation system does not. Example: the perceived bifurcation consists of a private (not encoded) road (caused by a difference in encoding, similar to failure mode 1a).
- Failure mode 2b: the system refers to a decision situation the driver has not perceived.

Spacial case: a complex situation is encoded as one decision situation by one participant (navigation system or driver) and is actually composed of two consecutive decisions; e.g., a an exit from an highway which imediately splits in the exit proper and a by-pass road - my navigation system instructs me as follows: "take the exit and then take the exit", not very natural, but at least clear.

8 Conclusions

Progress in understanding spatial behavior and linguistic expression of spatial situations may more likely result from analyzing dynamic situation than from further refinements of static spatial situations. Navigation and navigation instructions are an important and probably fruitful use case.

Coding the decision situation based on geometric identification of bifurcation is sufficient for shortest path algorithm to find the optimal path, but not sufficient to produce instructions human drivers can reliably follow. The visual interpretation of the perceived environment often leads to a different encoding as a the graph encoding of the street network.

The different modes that lead to failures in communicating instructions—and as a consequence, missing the optimal path—are the result of differences in classifying the environment differently by navigation system and driver. The classification may be different with respect to road classes or different with respect where decision situations occur and instructions are expected. As a general statement, the lack of correctly shared context leads to failure in following the optimal path. Automatic identification of decision situations from images analysis of dynamic scenes seems promising to produce instructions that drivers translate properly in actions.

References

Fictive Motion in Language and Perception, NATO ASI on Cognitive and Linguistic Aspects of Geographic Space, 1990.

Richard Bird and Oege de Moor. *Prentice Hall International Series in Computer Science*, chapter Algebra of Programming. Prentice Hall Europe, 1997.

E. W. Dijkstra. A note on two problems in connection with graphs. *Numerische Mathematik*, (1): 269–271, 1959.

Max J. Egenhofer and David M. Mark. *Spatial Information Theory - A Theoretical Basis for GIS*, volume 988 of *Lecture Notes in Computer Science*, chapter Naive Geography, pages 1–15. Springer-Verlag, 1995.

A. U. Frank. Gi can improve efficiency of economy by 15% - a pamphlet in support of gi2000. Technical report, 1998.

Andrew U. Frank. Qualitative spatial reasoning about distances and directions in geographic space. *Journal of Visual Languages and Computing*, 1992(3):343–371, 1992.

C. Freksa. *Cognitive and Linguistic Aspects of Geographic Space*, chapter Qualitative Spatial Reasoning, pages 361–372. NATO ASI Series D: Behavioural and Social Sciences. Kluwer Academic Press, 1991.

R. H. Güting and M. Schneider. *Moving Objects Databases*. Morgan Kaufmann, Elsevier, 2005.

Kathleen Hornsby and Max J. Egenhofer. *Spatial Information Theory - A Theoretical Basis for GIS (International Conference COSIT'97)*, volume 1329 of *Lecture Notes in Computer Science Vol.1329*, chapter Qualitative Representation of Change, pages 15–33. Springer-Verlag, 1997.

Torsten Hägerstrand. *Innovation Diffusion As a Spatial Proces*. University of Chicago Press, 1967.

Roberta L. Klatzky. *Spatial Cognition - An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, volume 1404/1998, chapter Allocentric and egocentric spatial representations:definitions, distinctions, and interconnections, pages 1–17. Springer, 1998.

Werner Kuhn. Image-schematic ontology integration. In *Spatial Representation and Reasoning in Language: Ontologies and Logics of Space*, 2010.

Benjamin Kuipers. *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge*. The MIT Press, 1994.

Stephen C. Levinson. *Language and Space*, chapter Frames of Reference and Molyneux's Question: Crosslinguistic Evidence, pages 109–170. MIT Press, 1996.

Damir Medak. Lifestyles - an algebraic approach to change in identity. In *Spatio-Temporal Database Management*, pages 19–38, 1999a.

Damir Medak. *Lifestyles - A Paradigm for the Description of Spatiotemporal Databases*. PhD thesis, Department of Geoinformation, 1999b.

Erwin Schröder. *Vorlesungen über die Algebra der Logik (Exakte Logik)*. Teubner, 1890.

M. White. *Geographical Information Systems: Principles and Applications*, volume 2, chapter Car Navigation Systems, pages 115–125. Longman Scientific & Technical, 1991.