Is AI Ready for Spatial Intelligence?

Stephan Winter

Department of Geomatics, The University of Melbourne

Abstract
Alan Turing famously approached intelligence of a machine by its communication behavior. A machine's spatial communication behavior would then define its spatial intelligence, in Turing's sense of successfully imitating human spatial communication behavior. This opinion and review paper asks whether the time has come for the spatial cognition and computation community to hold a grand challenge on spatial intelligence.

Keywords: spatial intelligence, spatial communication.

1 Introduction
Location is relevant everywhere: in information search (observing and understanding location) as well as in decision support (talking about location), and hence, in human-computer interaction (Dourish, 2001). With emerging ubiquitous computing by all sorts of small computing devices this prospect is impacting such diverse areas as spatial search, emergency call centers, command and control, navigation systems, ambient intelligence, location-based gaming, and physical coordination between mobile agents.

Location is an essential part of the context of human activity and decision making. And yet, the communication behavior of current services is far from the behavior of a human communication partner. The current service paradigm is still (a) reliance on a coordinate-based concept of space, (b) treating time as independent from space, and (c) lacking any other spatial semantics than what a gazetteer provides. In comparison, people memorize, reason about, and communicate about places and their relationships, prefer qualitative over quantitative descriptions, cope easily with incomplete and vague data, and consider intuitively the communication context in their interpretation or generation of location expressions.

Correspondence concerning this article should be addressed to Stephan Winter, Department of Geomatics, The University of Melbourne, VIC 3010, Australia; email: winter@unimelb.edu.au.
This paper asks the question whether the field is mature enough to bridge this gap between human communication behavior and current services’ communication behavior. As a community, spatial cognition and computation emerged over the last two decades (early milestones were Frank & Campari, 1993; Frank, Campari, & Formentini, 1992; Mark & Frank, 1991), and progress of the field is made even beyond this community, e.g., in natural language processing or human-computer interaction. Are we in a state to bring our collective and inter-disciplinary knowledge together to address a bigger issue: the dialog between people and the machine about all things spatial?

The paper culminates in suggesting a grand challenge, a restricted Turing test, as an inspirational vision or a benchmark for this community, or at least the part of this community that is working in the domain of cognitive ergonomics in human-computer interaction about geographic space. I admit that other aspects, for example, efficient machine reasoning and interaction are left out by this idea. However, efficiency was not in Turing’s mind when he thought about intelligence.

2 Spatial Intelligence

Turing, laying the ground for what later became known as artificial intelligence, starts his landmark paper *Computing Machinery and Intelligence* with the words: “I propose to consider the question, ‘Can machines think?’” (Turing, 1950, p. 1). Already Turing himself was aware of the hard problem of defining thinking or intelligence. To avoid the problem he came up with an elegant, operational suggestion: an anthropomorphic imitation game. In this game—nowadays called *Turing test*—a person is supposed to find out whether it is communicating (via teletype) with a machine or another person. Turing equaled anthropomorphic communication behavior with *being intelligent*. If the player cannot distinguish between machine and person the machine passes the Turing test.

However, his suggestion of the game sparked an ongoing controversy in artificial intelligence and beyond. This controversy entwines around the notion of thinking or intelligence: is imitating human communication behavior an indicator for intelligence? Searle’s Chinese Room experiment (Searle, 1980) is a prominent objection. The experiment basically says that a computer, programmed to understand Chinese, could also be replaced by a person running this computer program by hand: both do not understand Chinese. Yes, digital computers are mindless; they manipulate symbols in an order they were programmed. Accordingly, a program adds programmers’ knowledge to a computer, even an ability to learn and hence to act in ways not predictable by their programmers. This teaching of abilities and knowledge means a computer can only be appropriately programmed to pass the Turing test (what Searle calls *weak AI*), without a chance to claim having consciousness or a mind (what Searle calls *strong AI*). Obviously the Turing test relies only on the
communication behavior, i.e., the cognitive and linguistic performance capacities of a computer. It does not require to internally function like a human. Accordingly, we will abstain in the following from using the word *thinking*, and render our expectations more precisely to an imitation of a person’s communication behavior, calling this *intelligent* behavior.

In the current context we are seeking an operational definition for *spatial intelligence*, which means we take the liberty to restrict a Turing test to the limited scope of conversation about geographical space. A machine shows spatial intelligence if persons, initiating a dialog about geographic space, cannot find out from the communication behavior of the machine whether they are communicating with a navigation service or, through the machine, with another person—a spatially intelligent agent. Reflecting our initial interest, we further restrict the tenor of the conversation to satisfying a spatial information need, such as expressed in questions like “Where did it happen?” or “Can you tell me the way to the train station”. To be generous—extending Turing’s rules of communication—we might even allow for a graphical and gestural communication interface, since people enquire and describe geographic information graphically and by gestures as well. We call this imitation game the *spatial Turing test*.

### 3 Research towards Spatial Intelligent Behavior

Simple, everyday questions such as “Where did it happen?” or “Can you tell me the way to the train station” are answered with ease by a human informant familiar with the environment, but they are challenging a computing machine for multiple reasons, among them language processing, spatial semantics, and context, including the particular environment and the intentions and physical abilities of the enquirer (Figure 1).

![Google search interface](https://via.placeholder.com/150)

**Figure 1:** Spatial dialog is a challenge for multiple reasons, including spatial semantics (which features count as ‘train stations’) and context (the traveling mode of the enquirer, the supposed knowledge of the environment of the enquirer, the knowledge in the environment, the intentions of the enquirer to resolve ambiguities, etc.) (Interface: © Google, 2009)

This section provides a rather selective, but representative review of related recent research. It shows two properties of our current state of knowledge: (a) contributions are (necessarily) on narrow problems and not integrated yet, and (b) areas of good progress are seen alongside areas unattended.
3.1 Understanding Spatial Expressions

Understanding spatial expression is a sub-problem of natural language understanding (J. Allen, 1987). While this in general is a challenge even for a restricted Turing test, in our context spatial language is in the focus.

Two areas of research can be distinguished. Spatial linguistics (Levinson, 2003) starts from discourse analysis (e.g., Klein, 1979; Mark & Gould, 1995) and is interested in the understanding of larger structures of an expression, such as the use of spatial relations between individuals (e.g., Retz-Schmidt, 1987; Tenbrink, 2007), which are so fundamental for human experience that they are even used metaphorically (Lakoff, 1987). Other topics of relevance are context, in form of deixis (Klein, 1982), and perspective (e.g., Goschler, Andonova, & Ross, 2008).

The other area of research is less concerned with structure and relations, but with the meaning of individual terms, which can be either types in spatial taxonomies, or names of individuals. Georeferencing links between a set of well-known place names and their position in coordinate space, their taxonomic type according to a well-known taxonomy (Hill, 2006). Matching between or fusion of different taxonomies is studied (Bittner, 2007; Kokla & Kavouras, 2001; Riedemann & Kuhn, 1999) and helps in the process of making sense of a spatial expression. With respect to individuals, geographic information retrieval studies methods for the spatial interpretation of individual place names (Arampatzis et al., 2006; Jones, Purves, Clough, & Joho, 2008).

3.2 Spatial Reasoning

Searching for representations of qualitative spatial relations and for models of spatial reasoning in computers was inspired by linguistic and cognitive research (Talmy, 1983; Tversky & Lee, 1998). In their cognitive spatial representations and language people have a preference for qualitative over quantitative measures: qualitative distinctions are learned earlier (Piaget & Inhelder, 1956) and require less detail in representation and communication (Cohn & Hazarika, 2001; Frank, 1992, 1996; Kuipers, 1994). An intelligent service must understand qualitative spatial descriptions, and reflect (but not necessarily implement) a reasoning based on qualitative spatial relations.

Spatial reasoning has already some history in spatial cognition and computation. Besides of a longer tradition in logic (e.g., J. F. Allen, 1983; Kuipers, 1978) in this community the interest in qualitative relations probably started with Egenhofer’s dissertation (Egenhofer, 1989), which for the first time came up with the 4-intersection model to characterize topological relations. His model was quickly taken up by industry (Herring, 1987), and also tested for linguistic significance (Mark et al., 1995). Parallel to Egenhofer’s 4- and 9-intersection model an alternative model based on first order logic and the properties of connectedness (Clarke, 1981) was developed (Randell, Cui, & Cohn, 1992). Both models expanded into the relations between regions with
indeterminate boundaries, represented as rough sets (Clementini & Di Felice, 1996; Cohn & Gotts, 1996).

Other spatial relations were added: cardinal and relative directions, or orientation (Frank, 1991; Freksa, 1991, 1992), and distance (Frank, 1992; Hernández, Clementini, & Di Felice, 1995). Qualitative distance models are a special case in this list. Although algebras can be developed for distance relationships, they lack a grounding of their symbol set. Nearness is context-dependent (Worboys, 2001), in contrast to topology and orientation.

Also quantitative types of spatial reasoning impact on the acceptance of a reasoning result. Among them the most relevant in the given context is route choice. While it is not completely understood how people choose their route (but see Bovy & Stern, 1990; Timpf, Volta, Pollock, Frank, & Egenhofer, 1992; Wiener & Mallot, 2003), the industry observes that people often complain about automatically generated routes because these—typically the geometric shortest ones—can be relatively complex (Duckham & Kulik, 2003). Route choice models reflecting human travel behavior are a topic in operations research (e.g., Bekhor, Ben-Akiva, & Ramming, 2006; Frejinger, 2008; Hoogendoorn-Lanser, van Nes, & Bovy, 2005).

3.3 Generating Spatial Expressions

Results from a reasoning process have to be communicated in a way that imitates human communication behavior. Sometimes this goal is also labeled ‘cognitively ergonomic’ information provision (Klippel, Hansen, Richter, & Winter, 2008).

One of the earliest examples is probably the abandoning of traditional route maps in favor of sketches that follow cognitive principles in their selection and distortion (Agrawala & Stolte, 2000, 2001). Related, but not specific to a particular communication mode, is the systematic study of the principles of grouping larger chunks of elements together and reducing by this way the number of instructions within a route description (Klippel, Tappe, & Habel, 2003). An important role for chunking play (point-like, linear and areal) landmarks (Richter, 2007), as landmarks at route decision points or along long segments in general helps to enrich route directions. Consider the early implementations of route communication by landmarks in Figure 2: ideally landmark-based route directions can get rid of quantitative geometric distances, which better reflects the preference by people for providing qualitative descriptions (Denis, 1997; Denis, Michon, & Tom, 2007).

A prerequisite for using landmarks is research into the capture of salience of places, based on the human experience of space. Operational models to identify elements in a spatial database that are experienced as salient for a passerby help to fill spatial expressions. Work has been done to identify general landmarks (Elias, 2003; Raubal & Winter, 2002), or landmarks along a route (Klippel & Winter, 2005).
Another step forward is an operational model to generate destination descriptions: descriptions that do not describe how to find to a destination, as the traditional turn-by-turn directions do, but where this destination is (Richter, Tomko, & Winter, 2008; Tomko & Winter, 2009). This communication pattern can be frequently found between human communication partners if the informant can infer that the recipient has at least a coarse idea of the environment.

Destination descriptions already come up with a cognitively ergonomic principle, borrowing from the hierarchic organization of cognitive spatial representations (Couclelis, Golledge, Gale, & Tobler, 1987; Hirtle & Jonides, 1985): a variation of granularity within spatial descriptions. A need for this ability was confirmed recently (Tenbrink & Winter, 2009).

Where the previous models were mostly concerned with the selection of cognitive ergonomic references, the last step in the processing chain would be the generation of expressions in any language, e.g., verbally or graphically (e.g., Dale, Geldof, & Prost, 2005).

Figure 2: First web route planners providing guidance by ‘landmarks’. The identification, selection and integration of ‘landmarks’ in their directions is implemented to different levels of sophistication (© Yahoo!, AA, Sensis 2009).
4 A Grand Challenge in Spatial Intelligence

A *Grand Challenge in Spatial Intelligence* is a Turing test competition restricted in scope to the spatial domain of environmental and geographic scales (Montello, 1993), and restricted in tenor to seeking or providing spatial information. In the challenge, services have to demonstrate ‘human-like’ communication abilities about space and time, such that test players will blindly communicate with candidate services and other persons and have to decide in each case whether they are talking with a person or a computer.

The competition can be limited to specific interfaces (e.g., verbal or graphical) or to particular modes (e.g., under lab conditions or in-situ). Also, the design of the competition must provide for a sensitive issue: communication about spatial information is language and culture specific. While the choice of a single language may be necessary to enable judges to assess the performance of different services, a service could compete for an assessment within a particular cultural background.

Other disciplines have their own *Grand Challenges*. Among the thematically close ones are:

- The Loebner Prize\(^1\), a restricted general Turing test
- The Graph Drawing Contest\(^2\), in conjunction with the annual symposium on graph drawing
- The DARPA Urban Challenge\(^3\), an autonomous vehicle research and development program,
- Robocup\(^4\) (“By the year 2050, develop a team of fully autonomous humanoid robots that can win against the human world soccer champion team.”) and RobocupRescue\(^5\).

All of these challenges are addressing an academic community and are judged for scientific merits, in contrast to other challenges that are designed to popularize the use of specific products (e.g., the Google KML Challenge\(^6\), Teletas’ LBS Innovator Series\(^7\), or NAVTEQ’s Global LBS Challenge\(^8\)), and which are, accordingly, judged by representatives of the commercial sector.

The grand challenge in spatial intelligence is of the first kind. Competitors would be teams of researchers, and judges would be distinguished researchers without any conflict of interest. However, similarly to the DARPA Grand Challenge, the competing teams would rely on support and collaboration with

\(^1\) [http://www.loebner.net/Prizef/loebner-prize.html](http://www.loebner.net/Prizef/loebner-prize.html)
\(^4\) [http://www.robocup.org/](http://www.robocup.org/)
\(^5\) [http://www.robocuprescue.org/](http://www.robocuprescue.org/)
\(^6\) [http://www.google.com/educators/kml_contest.html](http://www.google.com/educators/kml_contest.html)
\(^8\) [http://developer.navteq.com/site/global/home/p_home.jsp](http://developer.navteq.com/site/global/home/p_home.jsp)
industry partners. What for the DARPA Grand Challenge are vehicles and electronics, is for the Grand Challenge in Spatial Intelligence an embodiment of the intelligent service (or the person behind) by a mobile device that serves as sensor and communication platform. Sensors would at least provide location, orientation and sound, but could include also imaging or distance sensors directed towards the environment or the user, for various purposes such as voice input or use as a pointing device. Wireless communication with a remote computer is essential: at this backend is either the intelligent service or a person. Similarly, the challenge needs the commitment of geographic data providers. Since spatial data is critical and can be quite diverse and hard to obtain, data should be provided quite in advance of the event to all competitors. Competitors need to tune their algorithms to the data provided, but they also may interpret the provided data for their own purposes, or add to that data. Other in-situ challenges, such as RobocupRescue, have their standard environments\(^9\); maybe the grand challenge in spatial intelligence should have a standard mobile device and a well-documented standard data format.

An impartial steering committee would set up a statute of the rules, oversee the organization, guarantee for the impeccable reputation, and look for the viability of the challenge. Ideally such a challenge has a prominent and potent industry sponsor with long-term commitment and no further interference in the process, and is held in conjunction with one of the major conferences to reduce the travel costs and efforts of all participants.

5 Summary

Alan Turing famously approached the problem of defining intelligence of a machine by its communication behavior. Spatial communication behavior would then define spatial intelligence, in Turing’s sense of successfully imitating human spatial communication behavior. This paper has proposed a Grand Challenge in Spatial Intelligence as a Turing test restricted in scope and tenor to spatial information.

So, is AI—or our inter-discipline of spatial cognition and computation—ready for a grand challenge in spatial intelligence? The discussion in this paper has revealed that a set of fragmented ideas, theories and models exist already, but also that no integrating, large-scale demonstrators (including the required data sets) exist. Commercialization is in its infancy (Figure 2); in lack of appropriate data only simplified versions of the current body of knowledge are implemented. In this situation, the Grand Challenge in Spatial Intelligence could come just in time to serve as inspirational goal and catalyst for an integrative type of research. Borrowing from others: By what year will we have developed a ubiquitously available service, a situated and embodied companion that acts and like a (situated, embodied) human informant?

\(^9\) [http://www.isd.mel.nist.gov/projects/USAR/]
Acknowledgements

This work is supported under the Australian Research Council’s Discovery Projects funding scheme (project number 0878119). Comments by Matt Duckham are acknowledged.

References


