Towards general spatial intelligence

How to cite:

For guidance on citations see FAQs

© 2008 The Authors
Version: Accepted Manuscript
Link(s) to article on publisher’s website:
http://en.scientificcommons.org/47680853

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Towards General Spatial Intelligence

Vlad Tanasescu and John Domingue

Abstract.

The goal of General Spatial Intelligence is to present a unified theory to support the various aspects of spatial experience, whether physical or cognitive. We acknowledge the fact that GIScience has to assume a particular worldview, resulting from specific positions regarding metaphysics, ontology, epistemology, mind, language, cognition and representation. Implicit positions regarding these domains may allow solutions to isolated problems but often hamper a more encompassing approach. We argue that explicitly defining a worldview allows the grounding and derivation of multi-modal models, establishing precise problems, allowing falsifiability. We present an example of such a theory founded on process metaphysics, where the ontological elements are called differences. We show that a worldview has implications regarding the nature of space and, in the case of the chosen metaphysical layer, favours a model of space as true spacet ime, i.e. four-dimensionality. Finally we illustrate the approach using a scenario from psychology and AI based planning.

1 INTRODUCTION

Is there a common denominator to the way we think about our spatial environment (representation), the way we find our way around (orientation, way-finding), our subjective experience of the external world (sense of place) or the way we coordinate our most common as well as more precise movements (mobility, manipulation)? We call any theory that attempts to unify these various experiences of space a theory of General Spatial Intelligence (GSI). By limiting intelligence to its spatial aspects we exclude non spatial activities such as writing a piece of music or reading a page of Tolstoy. However it is reasonable to think that a theory explaining spatial experience would fit nicely into any theory of general intelligence.

A basic use case that illustrates several “attitudes” toward space is an extended version of the “changing a light bulb” scenario. The original use case originates in the literature on ad-hoc categories [2] [3] whilst a modified version has recently been introduced in GISc1 [12]. It involves an intelligent agent having to change a light bulb in an office room using the various elements available. We slightly extend the use case by introducing the activation of a switch before realizing that the bulb is faulty (as an element of manipulation), as well as the need to go to another room to get another bulb (orientation).

We will demonstrate that even the cognitive operations involved in this scenario can be a part of a theory of GSI.

As “intelligent” natural agents, we do not only evolve in space and time but also reflect about them. To represent space, i.e. to communicate about it in a uniform way through formal (mathematic) or informal (natural language) means, we use distinct knowledge representation models. GIScience (or GISc) is the science of spatial representation [15]. To achieve geographic representation, and describe generic notions such as place [4] [14], or to handle precise problems such as land cover [1], GISc has witnessed the use of logic based ontologies, multi-representation models, image schemata, affordances or conceptual spaces. Moreover, the formalisms used range from mainstream programming languages and algebraic techniques, to functional or logical formalisms. Also, beside variations of models and formalisms at the representation level, spatial knowledge is acquired, experienced and shared at levels that can be, for example computational, cognitive, or supported by language.

These various dimensions of GIScience investigation consume theories – or “organized systems of accepted knowledge that apply in a variety of circumstances to explain a specific set of phenomena” – to produce implicit or explicit models, or “hypothetical [and simplified] descriptions of a complex entity or process”. Rather than questioning the validity of these approaches we believe that an attempt should be made to ground them in an underlying theory of General Spatial Intelligence. Indeed, without an underlying theory, the juxtaposition of models and domains to describe the experience of space gives the unsatisfactory appearance of a “collage”. Whether a unified theory is indeed possible is, of course, a research issue beyond the scope of this paper.

The term general in GSI does by no means mean universal, but simply underlines the aim of being as encompassing as possible. Indeed a theory of GSI needs to aim for universality in order not to exclude aspects of the GISc worldview from the start. Therefore GISc can be defined negatively as a) the rejection of any theory of space that describes only one aspect of spatial experience and b) the rejection of any theory that does not make all (or at least several) aspects of a worldview explicit. The advantages of a GSI theory are twofold:

1. Grounding: an underlying theory provides a justification for a particular model produced or used;
2. Fundational: an underlying theory, if deemed general, allows to develop new models based on the assumption that they should match the underlying.

To use a metaphor, an underlying theory can be compared to the bottom of a swimming pool, that can be used either to rest when tired of swimming or as a support to bounce without much effort toward new locations.

We will start by presenting Raper’s account of worldviews in GISc, and augment it with the notion of poles. We will then provide a description of the representation layer and suggest that GISc cannot be achieved at this layer only. We will then make the metaphysical layer explicit, using an instance of process philosophy. The

1 Knowledge Media Institute (KMi), The Open University, United Kingdom
2 in reference to Artificial General Intelligence [10], of which it may be considered as a subset.
3 we will use both GIScience and GISc, the former can be considered as an abbreviation of the latter.
4 http://wordnet.princeton.edu/perl/webwn?l=theory
5 http://wordnet.princeton.edu/perl/webwn?l=model
The notion of poles will be used to describe and apply constraints on the notions of space, mind, language and agency. Finally we will apply the resulting theory to a version of the changing a light bulb use case.

2 WORLDVIEWS

As Thomas Kuhn quoted by Raper describes it “a worldview can be seen as a set of assumptions and commitments to which a ‘research programme’ subscribes” [15] (p. 5). Still according to Raper [15], a worldview is composed of at least these fundamental issues:

- **Metaphysics**: the theoretical grounding and conceptualisation of the notion of ‘world’ employed;
- **Ontology**: the approaches by which the contents of the conceptualisation of the world are defined, ordered and signified;
- **Epistemology**: the procedures by which knowledge of the conceptualised world is established and evaluated;
- **Philosophy of Mind**: the nature of human knowledge of the conceptualised world;
- **Linguistics**: the nature of language and its role in communication and the construction of meaning;
- **Cognitive Science**: theories and models of cognition, including perception, the nature of intelligence and the functioning of the mind; and
- **Representation**: the nature of computation employing symbolic and informational representations of human knowledge.

Most investigations in the spatial domain leave one or several of these aspects implicit, providing theories for only a few of them, without acknowledging the importance of the others. This situation is not restricted to GIScience but we believe that making explicit statements about these issues would provide multiple advantages, notably:

- **Completeness**: the aim of a theory of GSI would be to be comprehensive, at least teleologically, i.e. aim to eventually explain all of spatial experience, subjective and objective;
- **Incremental Progress**: with a theory aiming for completeness, the results to specific problems could be tied to the underlying theory and hence reflected in other models;
- **Falsifiability**: one would be able to define precise problems against which the theory could be tested;
- **Modelisation**: a well defined theory would allow derivative models, i.e. simplifications of subsets of a theory, without giving the impression of a collage;
- **Philosophical Jackpot**: finally, the possibility exists that such a structure of interdependent elements, modified and completed by new problems, achieves some new representational power, allowing for example embodied mechanical agents to reach a level of agency that could be qualified as Artificial General Spatial Intelligence.

Elements of a worldview follow a layered structure, rooted in metaphysics and ontology and leading to representation through the layers of epistemology, language, mind, etc. [15]. However, in a theory of General Spatial Intelligence, not only should worldview layers interact with each other but they are also required to be mutually consistent and reducible to each other. Indeed a successful description of the sense of place for example, seems very unlikely without any reference to mind, language and cognition. In the same way a theory of mobility that does not corroborate physical laws would be easily falsified. Moreover, one should be able to describe KR models in terms of language or cognitive features, in the same way as, for example, a high level programming language can be expressed in low level binary operations. Notably, as we aim for completeness, we should strive to establish a link between the two poles constituted by the representation layer and the metaphysical one. These links, when established will contribute to the orientation of all other layers, similarly to the poles of a magnet with iron filings but in both directions (cf. Figure 1), even if layers are not absolutely equivalent and that the transitions may loose some specificity.

![Figure 1. Worldview and Worldview Poles](image)

The truth of statements about any of the layers is not a concern here, as truth itself is considered to be a part of the worldview (originating in the epistemological layer). This approach could be qualified of postmodern, as it assumes that science, being the conjunction of knowledge and truth, may be, at least in parts, dependent of the social and historical processes in which it evolves. However we believe that the advantages of a theory of GSI can satisfy non postmodern adopters as well, who should feel free to propose a worldview satisfying the constraints of GSI and see it true and absolute, as long as it provides the same explanatory power.

3 A WORLDVIEW POLES FOR GSI

Representation of spatial experience, the aim of GIScience, and one of the poles we identified in a worldview, is confronted to issues regarding what Knowledge Representation (KR) model is appropriate to a given situation. Whereas any representation is ungrounded without an underlying layer, which can be ultimately grounded by metaphysics.

3.1 Representation: at Worldview’s End

In order to describe different domains, various KR models has been advocated in GIScience. To our knowledge, these models all fit in the categories of logic based ontologies, affordances, image schemata, conceptual spaces and multi representation. Each of these models have specific components that are used to describe aspects of spatial knowledge:

- **Logic based ontologies** describe a domain using **categories** and by defining relations between them. Categories have **attributes** and are related by **is-a relationships** to form taxonomies. **Instances**

Or, as Gilles Deleuze puts it, “Philosophy can speak of science only as of a cloud.” [5]
are individuals belonging to a category. Moreover, some ontology languages provide procedures, functions, axioms and (production) rules.

- **Multi representation** models promote the notion of context-dependent identity. Instead of linking an entity to a unique set of categories with fixed sets of attributes, multiple representations are provided depending on the context. These representations are not only visual (e.g. an airport being represented geometrically as a point at world scale and a complex structure at a closer scale) but also concerns the attributes, which can change depending on the particular representation dictated by the context [24].

- **Conceptual spaces**, as introduced by Gardenfors [8], present individuals as points in geometric multi dimensional quality spaces. As an example, a particular colour is represented as a point in the three dimensions of the primary colours. This representation allows to easily measure similarity between two individuals as the distance between two individual points.

- **Image schemata** are described as “a small number of parts and relations, by virtue of which it can structure indefinitely many perceptions, images, and events” [13], and are basically a set of structures which support relations. Examples are container, balance, compulsion, blockage, counterforce, etc. Schemas can be combined to create complex structures. For example a crossroad is composed of two paths, which are themselves based on the image schemata link and support.

- **Affordances** are the constitutive concept of ecological psychology [9]: they are possibilities for action latent in an environment. As actions have to be carried by someone, the users’ capabilities have an important role in the definition of an affordance (e.g. a wall does not afford to be climbed if I am not tall or fit enough or do not have a ladder).

If these paradigms seem to be superficially compatible (e.g. quale in DOLCE [7] are inspired by conceptual spaces) only minimal parts of each model can be mapped to another one and integrations tend to disregard the specificities of the paradigm (Gibson’s affordances for example, are often assimilated to functions, or their meaning is disregarded). Therefore KR models are often used in isolation (e.g conceptual spaces in [16] or in [1]) which reveals their expressive limits. As an example, it is indeed quite difficult to emulate is-a hierarchies with conceptual spaces. Moreover, some statements are simply not understandable without an underlying theoretical layer: for example direct perception in ecological psychology, the theory that affordances are perceived without the mediation of reasoning.

Incompatibilities between KR models are summarized in Figure 2. Hence we believe these incompatibilities between representation models are fundamental and cannot be solved at the representation layer. Each KR model expresses different aspects of underlying layers, in which radically different elements are at work: indeed affordances are not functions, and classes/categories are not prototypes, although they may provide approximations for each other.

Since KR models are essentially irreconcilable, there is a need to establish a grounding at a lower level. As our aim is GSI, i.e. trying to explain spatial cognitive and language related attitudes as well, we will make an attempt to find this foundation at the lower possible layer, i.e. the opposite pole, by introducing a metaphysics able to explain these differences.

---

![Figure 2. Knowledge Representation in GIScience](image)

### 4 Metaphysics For GSI

Metaphysics seeks to answer questions about the ultimate nature of reality. Classical KR has been grounded on an (often implicit) metaphysical framework that places emphasis on things and substance. Such a metaphysical layer favours KR models based on identity, categories and attributes [18]. However, to ground a theory of GSI, the metaphysics we choose should provide a grounding for the different KR models, as well as guidelines for the other layers constituting the worldview. Given the nature of geographic space any metaphysics that aims to be relevant to GSI should at least display the following characteristics:

- **Granularity**: ability to explain the diversity of the KR models used in GISc;
- **Process Friendly**: many geographical objects, for example those represented as fields (e.g. the weather), are procedural by nature, and therefore a GSI metaphysics should be able to account for processes; and
- **Faultability**: issues of vagueness remind us that the conceptualization of geographic space is somehow subjective, and therefore the metaphysical should explain this difficulty.

In the next section we propose an instance of process metaphysics displaying these characteristics.

#### 4.1 Process Metaphysics and Differences

Process metaphysics identifies metaphysical reality with change and dynamism. As opposed to philosophies that posit an eternal truth at their core, process metaphysics defines the ultimate nature of reality as transformation and change itself. At a basic level, process metaphysics states that everything is a process, in movement, and that what we recognize as stable, only displays the relative stability of processes continuously acting to ‘keep it together’. Hence if at the ultimate level everything is a process, the reality is indistinct, pure process flux, pure activity. Variants of this have been advocated from Heraclitus of Ephesus (b. ca. 540 B.C.) teaching that panta rhei (“everything flows”) to Friederich Nietzsche, Henri Bergson, Alfred N. Whitehead and Gilles Deleuze.

In previous work, using only the notion of process, we have introduced differences as the ultimate element, not of reality (as process are in process metaphysics), but of meaning [19]. Indeed, positing
that everything is a process implies that processes process other processes, and by doing so, as a side effect, that some processes become isolated from the undifferenciated flux of activity that constitutes reality, starting to “make sense”. We call these processes differences. A difference is any process isolated by another process: it has a meaning by itself, but, being a process, can also be reduced to the differences/processes they are linked to. For example particular colours, shapes, a distinct word, a sound, an action, an event, a town, exhibit meanings I can grasp immediately if I have access to the appropriate processes to understand them, i.e. if these particular processes are part of the process that constitutes my individuality. However they can still be reduced to other differences that in some contexts may be called parts (e.g. for a “building”) or relations (e.g. for a “word”). Therefore differences are anything that makes sense by being isolated from the background by an actual process. Unlike symbols, differences are grounded as they cannot be dissociated from the processes that allow them to signify. Symbols become differences as soon they are recognized as meaningful.

4.2 Language and Representation

Differences produce meaning, without consideration to any particular KR. This is reminiscent of collaborative tagging in Web 2.0 applications, in which users are given absolute freedom in the elicitation of a label, or tag, independently of any pre-established categorisation, or indeed preexisting KR model. In this sense, collaborative tagging systems are expressions of pure meaning, as differences. Extreme Tagging Systems (ETS) are an attempt to generalize the use of collaborative tagging systems, to match the theory of differences [21]. Indeed, in ETS tags themselves can be tagged, as well as the relations between them. For example, if a representation of a car is tagged with “car”, “car” itself can be tagged with “wheels” and the relationship between the two with, for example, the tag “part-of”. These operations can be done collaboratively by multiple users on the same tag network, or individually on one’s own tag network. In brief ETS are semantic networks with no particular constraints on the vocabulary during the creation phase. However control mechanisms can be applied to the tagging operation itself, in the form of annotations (ratings, descriptions, etc). We have shown elsewhere [20] that ETS can be mapped to KR models, using for example a lexicon to find categories and affordances in the tags, in order to define the notion of place. Therefore we have a means to reach the representation layer from the metaphysical one. However, we are still lacking essential notions of GISc, such as space, and agency.

4.3 Space and Agency

The definition of space provides a perfect example of the application of the interaction between representation and metaphysics (illustrated in Figure 1). Indeed, in order establish a theory of GSI we should ponder if there is a need to introduce an independent concept of space, or if the underlying metaphysics already provide some useful abstractions that may support spatial representations. Moreover, for GSI, any notion of space should be linked to a notion of agency, i.e. become an environment for an agent. In this case, is there a need to introduce the complex collage usually characteristic of descriptions of agency (and start articulating concepts such as “plans”, “beliefs”, “goals”, etc.)?

If we answer the first question by stating that space should indeed be derived from the metaphysical layer, and allow differences at the representation layer, there are now two constraints that it should follow:

1. Procedural aspect: a metaphysic based on processes strongly suggests that time should be integrated with space in order to allow the natural execution of processes.

2. Differential aspect: we conceive differences between spaces according to the context, and therefore a representation of space should naturally adapt to these differences.

We believe that process metaphysics provides sufficient grounding for a notion of space supporting agency in a natural way. Indeed so-called “true spacetime” [6], i.e. an eternalist and perdurantist view of space [17], allows the representation of processes as timespace worms, i.e. 4-dimensional regions of spacetime. Following this view, a process such as an agent’s existence is defined as a region in spacetime, with its future and past are already given, and with a shape defined by its activity, or ‘movement’. This notion allows to solve many conundrums related to causality and agency. Indeed, causality becomes a consequence of the way that processes are perceived, i.e. differences, as in an undifferentiated view of spacetime both cause and effect participate to the same process. By attributing “meaning” to things, by differentiating them from the flux of becoming, we posit that difference A is cause of difference B. Another convenient feature of four-dimensionalism, regarding agency, is that future states of the world (the foundation of precision or planning) as well as past ones (the foundation of knowledge or memory) can be thought of in analogy with spatial regions. Indeed we can extend a theory of perception (the way we perceive our surrounding environment, a region of spacetime, for example visually) to cognitive operations in which an agent remembers its past (perception of a region of spacetime) or plans its future (perception of another region of spacetime). As such planning actions using past experiences in order to realize a foreseeable’ future corresponds to perceiving (possibly wrongly) regions of space time separated in time, i.e. a model of knowledge based planning is looking up toward the end of the staircase while climbing stairs, and at the same time using each stair as a support to climb higher. Models of perception can therefore be used to derive concepts such as “plans”, “desires”, “beliefs” of “goals”.

The second constraint imposed to spacetime is differentiation: we should be able to consider different notions of space according to the context. Indeed, when it comes to our experience of space, when moving, manipulating, or reading a map, space takes very different shapes and topologies. From the network geometry of paths in a urban environment, organized in zones and directions, to the 3 dimensional spaces of buildings and individual homes, divided by walls, connected by doors, extended by staircases and open to gardens, as well as the smaller environment of manipulable objects, all involve distinct types of spaces, having different topologies and geometries, as respectively the urbanist, the architect, and the robot engineer know. However, this differentiation should not affect the procedural aspect of space, as spacetime, that we already established.

Interestingly the need for different topologies and geometries has been acknowledged long ago in physical sciences. Indeed in modern physics shape is playing an essential role and older physical abstractions related to agency, such as forces, can be reduced to it. For example gravitational forces in the theory of general relativity are described as the action of the curvature of spacetime due to mass/energy fields [11]. Similarly any physical interaction seems to allow to be described in terms of curved multi-dimensional manifolds [22]. Moreover, as the geometry of space has become more
interesting than the absolute position of elements in it, space-time is often represented by a differential operator, the line element, which, independently of the coordinate system, describes the shape of a given space. For example the geometry of a 2 dimensional manifold shaped as a sphere is described by the differential element: 

\[ ds^2 = d\theta^2 + \sin^2\theta d\phi^2. \]

Here expressed in polar coordinates. In the same way “All geometry can be reduced to relations between distances; all distances can be reduced to integrals of distances between nearby points” [11] (p. 23).

We extend to agency the consequences of this procedural and differential description of space. Indeed four-dimensionalism as expressed here provides a surprisingly good account of direct perception: in the flux of processes constituting spacetime, direct perception of affordances is due to the fact that affordances are immediate result of our actions, and are therefore immediately (pre)visible. According to this theory the doorknob that affords turning is the immediate, i.e. differential in the sense of a function derivative, link to the foreseeable future of the door being opened, which coexists with the process that constitutes the agent’s activity (and existence). It needs to become meaningful, i.e. be isolated as a difference, in order to actualise this state of the world. Also the doorknob shapes the surrounding space for a particular agent by its affordances, allowing only a finite set of relevant movements to the foreseeable regions of spacetime it provides access to. Perception can therefore only be direct because it results movement, in a space presenting a given geometry allowing the affordance, a movement not only in space-time but in spacetime, that is oriented toward the completion of a process. Movement is constrained by the particular shape of the environment, and movement unifies both physical and cognitive operations. Similar conclusion has also been reached in [23], starting from the premise that the apparent intentionality of an agent can be explained by it following environmental gradients.

We will now informally describe a use case and show how such a theory of GSI can be applied to it.

5 CHANGING A LIGHT BULB

The process in which the agent is involved includes using an office room. The spacetime region occupied by this process, as perceived by the agent, involves switching on the light in the office, which involves movement from the three dimensional flat space of entering the room to the differently shaped and table-top like space of manipulating the switch.

After acting on the switch toward the expected spacetime region of bringing light to the room, and seeing that it does not occur, the agent’s perception is reoriented towards a different process in spacetime, in which the process of using the office to do some work involves the process of changing the light bulb.

Getting a new light bulb involves mobility in the spacetime defined by the compartimented 2D of the house. In this spatial context the agent orients itself in spacetime towards the expected spacetime region occupied by the new light bulb. In this case a drawer in the kitchen. Opening the drawer again involves a different space than the one followed to reach it.

Back in the office room the process of reaching the lamp involves the process of getting higher which involves using one of the present elements, traditionally a desk, a chair, or a pile-of-books.

Sliding the desk to climb on it leads the agent to forsece a spacetime region in which there are scratches on the floor, and this spacetime region leads to another in which blame can be felt. This acts as a negative affordance, and leads to other choices for reaching the bulb.

Eventually pushing and standing on the chair appears, involving its own spaces (flat 2D for the push operation, 3D for standing, etc.), while effectively changing the bulb involves a rotation with several degrees of freedom.

6 CONCLUSIONS

We have shown the need of a theory of General Spatial Intelligence, and hinted towards one in which agents evolve and act in their spacetime environment according to the shape of differential spaces induced by affordances. Future work will involve building the computational tools necessary to model spatial interactions.

REFERENCES