Information Processes Produce Imperfections in Data—How Does Information Infrastructure Compensate for Them?

Andrew U. Frank
Department of Geoinformation and Cartography
Technical University Vienna
Gußhausstraße 27-29/E127
A-1040 Vienna, Austria

V5 —SVN: 435 – 6229 Words

Abstract: Data quality descriptions consider the imperfections found in geographic data. These imperfections are caused by imperfect realizations of the processes that are used to collect, translate, and classify the data. The tiered ontology gives a sensible framework to analyze the data processes and the imperfections they introduce. Decision methods using the data are adapted to some of the imperfections and compensate for them. Additional methods to reduce negative effects of imperfections in the data on decisions are used when necessary.

1 Introduction

Geographic information is used for many different applications and therefore the assessment of the quality is increasingly of concern. A general thread of the discussion assumes that low quality is negative and focuses on methods to improve the quality. In this article I will demonstrate that the common assumption is wrong and highest quality is generally undesirable; only the necessary quality for a decision is useful and more precision is a waste of resources and more detail likely adds to confusion. Ordinary processing evolved to respect these limits; over time decision processes adapt to limitations of data collection and introduce compensations for these limitations.

I give first an ontology based treatment of *imperfections* in data and then identify methods the geographic information infrastructure uses to reduce negative effects of imperfections; these methods will be called *compensations*. The tiered ontology for geographic data (Frank 2001; Frank 2003) is used and extended to include the processes that are used to transform information between the tiers. The imperfections introduced by each process are assessed.

Human decision making is based on heuristics; a model of rationality would amount to perfect and complete knowledge—something humans cannot achieve. Even bounded rationality (Simon 1956) in an extreme interpretation is not a realistic model as it implies a rational decision when more information is necessary. A realistic, ecological model of human decision making takes into account that humans have limited computational resources and must make decisions in limited time (Gigerenzer et al. 1999). The focus of this paper is on producing a realistic ontology in the sense of *ecological rationality*.

In this article I build on the foundation of previous publications on ontology (Frank to appear 2008) and explore how imperfections in data processing are compensated. From a systematic review of the information processes follows what kind of imperfections these processes introduce and indicates how compensation methods can be used to reduce negative influences on decisions.

The novel contribution of the article is, firstly, the generalization of information processes to focus on the imperfections they produce and, secondly, the compensatory methods that are related to each kind of imperfection based on principles of ecological reasoning.

This paper starts in section 2 with a short review of a simplified tiered ontology. Section 3 shows how imperfections are introduced by data processing. Section 4 looks at decisions and how imperfections affect them. Section 5 then lists some strategies to compensate for imperfections of geographic data is used interoperability in a spatial data.

2 Ontology

An ontology describes the conceptualization of the world used in a particular context (Guarino et al. 2000): different applications may use different conceptualizations. A car navigation system determines the optimal path using the conceptualization of the street network as a graph of edges and nodes, whereas an urban planning application conceptualizes the same space as regions with properties. The ontology clarifies these concepts and communicates the semantics intended by data collectors and data managers, to persons making decisions with the data.

If an ontology for an information system contributes to the assessment of the usability of the data, it must not only conceptualize the objects and processes in reality but must also describe the information processes that link reality to the different conceptualizations. If an ontology divides conceptualization of reality in tiers, e.g. (Frank 2001; Smith et al. 2004), then it must describe the processes that transform data between tiers.

2.1 Tier O: Physical Reality

Tier O of the ontology is the physical reality, that “what is”, independent of human interaction with it. Tier O is the Ontology proper in the philosophical sense (Husserl 1900/01; Heidegger 1927; reprint 1993; Sartre 1943; translated reprint 1993); sometimes Ontology in this sense is capitalized and it is never used in a plural form. In contrast, the ontologies for information systems are written with a lower case o. The observed interactions between humans is only possible if we assume that there is only one, shared physical reality.

2.2 Tier 1: Observations

Reality is observable by humans and other cognitive agents (robots, animals). Physical observation mechanisms produce data values from the properties found at a point in space and time.

$v=p(x, t)$

A value v is the result of an observation process p of physical reality found at point x and time t. Tier 1 consists of the data resulting from observations at specific locations and times (termed point observation); philosophers sometimes speak of ‘sense data’. In GIS such observations are, for example, realized as raster data resulting from remote sensing (Tomlin 1983), similarly our retina performs many such observations in parallel.

2.3 Tier 2: Objects

The second tier of the ontology contains the description of the world in terms of physical objects. An object representation is more compact, especially if the subdivision of the world into objects is such that most properties of the objects remain invariant in time (McCarthy et al. 1969). For example, most properties such as color, size, and form of a taxi cab remain the same for hours, days, or even longer. They need not be observed and processed repeatedly. Only location and occupancy of the taxi cab change often and must be regularly observed.

The formation of objects—what Zadeh calls granulation (Zadeh 2002)—first determines the boundaries of objects and then summarizes some properties for the delimited regions before a mental classification is performed. For objects on a table top (Figure 1) a single process of object formation dominates: we form spatially cohesive solids, which move as a single piece: a cup, a saucer, and a spoon.

Figure 1: Simple physical objects on a table top: cup, saucer, spoon

Geographic space does not lead itself to such a single, dominant, subdivision. Watersheds, but also areas above some height above sea level or regions of uniform soil, uniform land management, etc. can be identified (Couclelis 1992). However, they are delimited by different properties and can overlap (Figure 2).

Figure 2: Fields in a valley: multiple overlap subdivisions in objects are possible.

2.4 Tier 3: Constructions

Tier 3 consists of constructs combining and relating physical objects. These constructs are generally socially coordinated. A physical object X is used to mean the socially constructed object Y in the context Z. For example, a special kind of stone in the ground counts as boundary maker in the legal system of Switzerland.

“X counts as Y in context Z” (Searle 1995, 28)

Social constructions relate physical objects or processes to abstract objects or processes. Constructed objects can be constructed from other constructed objects, but all constructed objects are eventually grounded in physical objects. The physical object can be a physical object in a situation like the cup in Figure 1 or a sign, which relates to a constructed object; e.g., the written or spoken word “cup” on the menu of a restaurant.

3 Information Processes Transform between Tiers

Information processes transform information obtained at a lower tier to a higher tier (Figure 3):

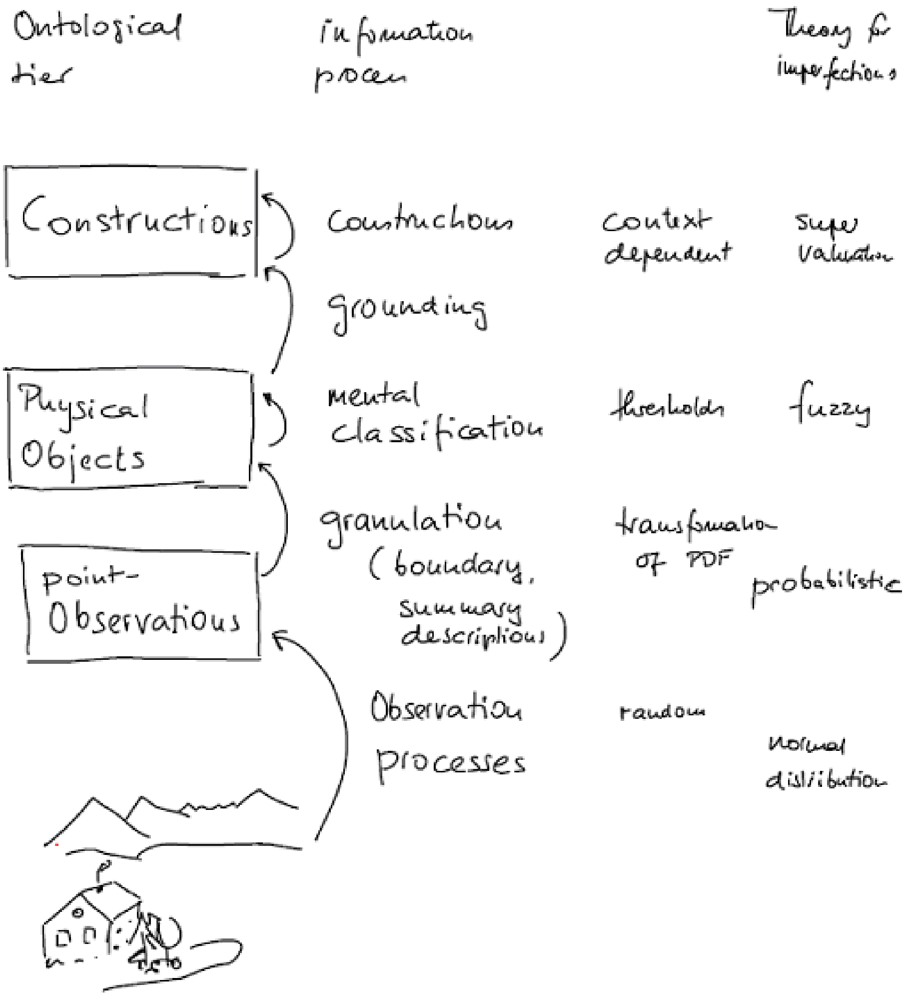

Figure 3: Tiers of ontology and information processes transforming data between them

All human knowledge is directly or indirectly the result of observations, transformed in often long and complex chains of information processes. All imperfections in data must be the result of some aspect of an information process (Figure 3). As a consequence, all theory of data quality and error modeling has to be related to empirically justified properties of the information processes. The production of complex theory for managing error in data without empirical grounding in properties of information processes seems to be a futile academic exercise.

The information processes will be analyzed in the following sections to understand their effects on data, specifically how they contribute to imperfections in the data.

3.1 Observations of Physical Properties at Points

The observations of physical properties at a specific point is a physical process that links tier O to tier 1; the realization of which is imperfect in 3 ways

- systematic bias in the transformation of intensity of a property into a quantitative (numerical) value,
- unpredictable disturbance in the value produced, and
- observations focus not at a point but over an extended area.

The systematic bias can be included in the model of the sensor and be corrected by a function. The unpredictable disturbance is typically modeled by a probability distribution. For most sensor a normal (Gaussian) probability distribution function (PDF) is an appropriate choice.

A sensor cannot realize a perfect observation at a perfect point in space or time. Any finite physical observation integrates a physical process over a finite region during a finite time. The time and region over which the integration is performed can be made very small (e.g., a pixel sensor in a camera has a size of 5/1000 mm and integrates (counts) the photons arriving in this region for as little as 1/5000 sec) but it is always of finite size and duration. Note that the size of the area and the duration influences the result (Openshaw et al. 1991).

The necessary finiteness of the sensor introduces an unavoidable scale element in the observations. The sensor can be modeled as a convolution with a Gaussian of the physical reality. Scale effects are not yet well understood, despite many years of being listed as one of the most important research problems (Abler 1987; NCGIA 1989b; NCGIA 1989a; Goodchild et al. 1999).

3.2 Object Formation (Granulation)

Human cognition focuses on objects and object properties. We are not aware that our eyes, but also other sensors in and at the surface of our body, report point observations, e.g., the individual sensors in the eye’s retina give a pixel-like observation, but the eyes seem to report about size, color, and location of objects around us. The object properties are immediately available, converted from point observations to object data without the person being conscious about the processes involved. Processes of object mental formation are found not only in humans, higher animals form mental representations of objects as well. Object formation increases the imperfection of data—instead of having detailed knowledge about each individual pixel only a summary description (summary value) of, for example, the middle wheat field in Figure 2 is retained. The very substantial reduction in size of the data is achieved with an increase in imperfection. The compact representation as a region requires few points for the boundary and achieves $1:10^5$ compression; it is a very powerful heuristics!

Object formation consists of two information processes

- boundary identification
- computing summary descriptions,

Mental classification is addressed in the next section.

3.2.1 Boundary identification

Objects are—generally speaking—regions in 2D or 3D that are uniform in some aspect. The field in Figure 2 is uniform in its color, tabletop objects in Figure 1 are uniform in the material coherence and in their movement: each point of the rigid object moves with a corresponding movement vector.

An object boundary is determined by first selecting a property and a property value that should be uniform across the object. It produces a region of uniform values and boundaries for these regions. Two different methods for determining object boundaries are:

A) By thresholds on the values *v* of interest: the object is the connected region of all point observations for which the value *v* is between a lower and an upper limit

B) By maximal change: the object boundary is where the value of the value *v* of interest changes maximally (Burrough 1996).

The location of the boundary derived by these two methods is not the same!

Assuming a PDF for the determination of the property of interest one can describe the PDF for the boundary line. The information process has an associated transformation function that transforms the PDF of the point observation in a PDF for the boundary line. (Figure 4)

3.2.2 Determination of descriptive summary data

Descriptive values summarize the properties of the object determined by a boundary. The value is typically an integral or similar summary function that determines the sum, maximum, minimum, or average over the region, e.g., total weight of a movable object, amount of rainfall on a watershed, maximum height in country (Tomlin 1983; Egenhofer et al. 1986).

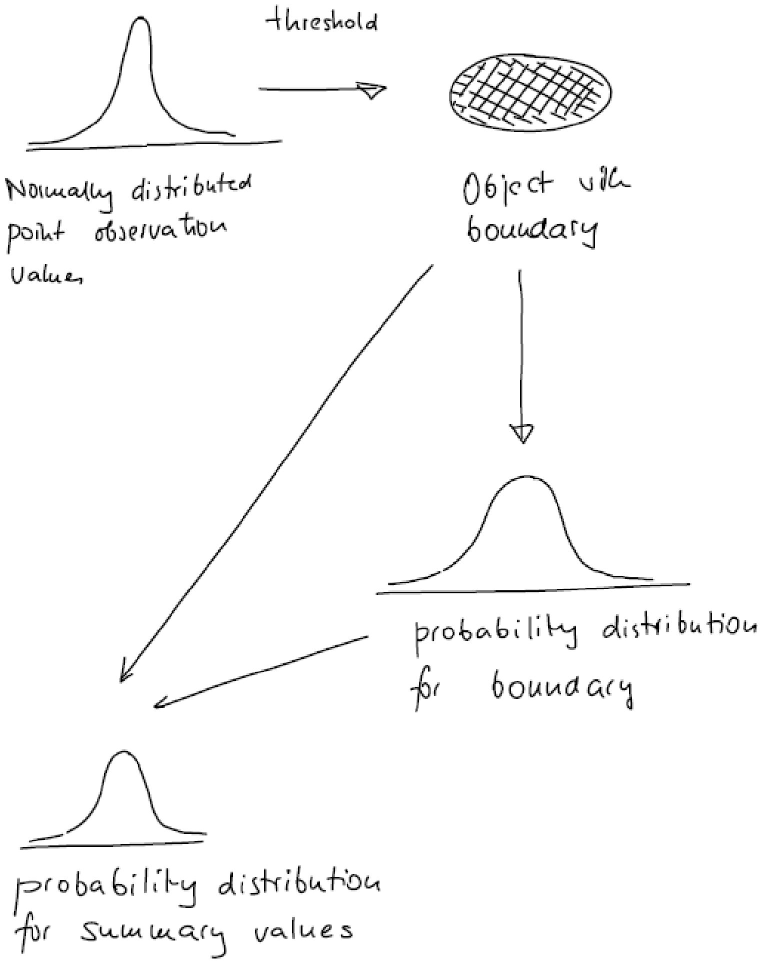

Figure 4: Transformation of probability, distinction functions from observations to boundary and summary value

If the observation information processes allow a probabilistic description of the imperfections of the values, than the imperfections in the object boundary and summary value are equally describable by a probability distribution. Given the PDF for the value of interest of the summary and the PDF for the boundary, a PDF for the summary values is obtained by transformation of the input PDF (Figure 5). It is an interesting question whether the PDF transformation functions associated with boundary derivation and derivation of summary values preserve normal distribution.

3.2.3 Mental classification

Objects once identified are mentally classified. On the tabletop, we see glasses, forks, and plates; in a landscape forest, fields, and lakes are identified. Mental classification is an information process internal to tier 2 related to “affordance” for the potential use of an object (Gibson 1986; Raubal 2002). Mental classification relates the objects identified by granulation processes to operations, i.e., interactions of the cognitive agent with the world. To perform an action, e.g., to dissolve sugar in coffee (Figure 1) requires a number of properties of the objects involved: cup must be container, i.e., having the affordance to contain a liquid, the object must be a liquid, etc.

I have used the term distinction for the differentiation between objects that fulfill a condition and those that do not (Frank 2006). Distinctions are partially ordered: a distinction can be finer than another one (e.g., drinkable is a subtaxon of liquid), distinctions form a taxonomic lattice (Frank 2006). The mental taxonomy adapts in the level of detail to the *situation* and can be much finer if the situation requires it than the one implied in the vocabulary (Ganter et al. 2005). Affordances (Gibson 1986) are in this view bundles of distinctions.

Humans classify unconsciously and immediately the objects we encounter and retain only the classification without verbal labels. Grouping of distinctions required for typical interactions form abbreviations. For example: the flat things that can be cut by a pair of scissors (i.e., paper), or the self-powered, movable things steered by a human passenger (i.e., cars). The classification in the mental taxonomic lattice is an abstraction reducing the amount of detailed information initially perceived in preparation for a probable decision. Instead of retaining detailed values for the decisive properties till the time of decision making only the classification is retained.

This abstraction process is cognitively plausible and supported by empirical evidence. If you interact with a household object (e.g., eat from a plate in a restaurant) and are later asked about detailed properties of the object you most likely realize that the properties you considered to classify the object as a plate were not retained, only the final classification (Randow 1992). The situation influences the interactions with the objects an agent considers; the relevant interaction determines which properties to use for object formation. All of this is summarized in a classification.

Distinctions reflect the limits in the property values of an object, where the object can or cannot be used for a specific interaction. The decision whether the values for an object are inside the limits or not is more or less sharp and the cutoff gradual (Figure 6). The distinctions and classifications are therefore fuzzy values, i.e., membership functions as originally defined by Zadeh (1974) (Figure 5).

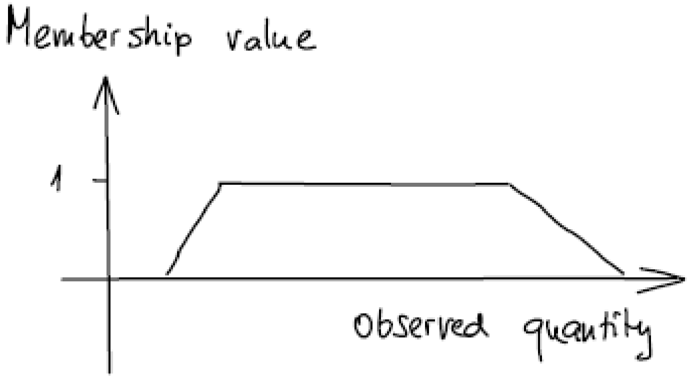

Figure 5: Classification of objects result in fuzzy membership values

3.3 Constructions

Constructions are concepts that are (1) mental units, which (2) have external representations (signs, e.g., words), (3) can be communicated between cognitive agents, and (4) are, within a context, without imperfection. The realm of constructions is linked through granulation and mental classification to the physical reality of physical objects and operations.

The agent’s direct sensory experience of the world is reflected in the agent’s experience of the world, an externally representable information image of reality is created duplicating the sensory “reality” in the brain. I call the constructions that stand for direct experiential reality grounding items. The classified sensory experience and the grounding items are isomorphic and are not consciously separable (Figure 6).

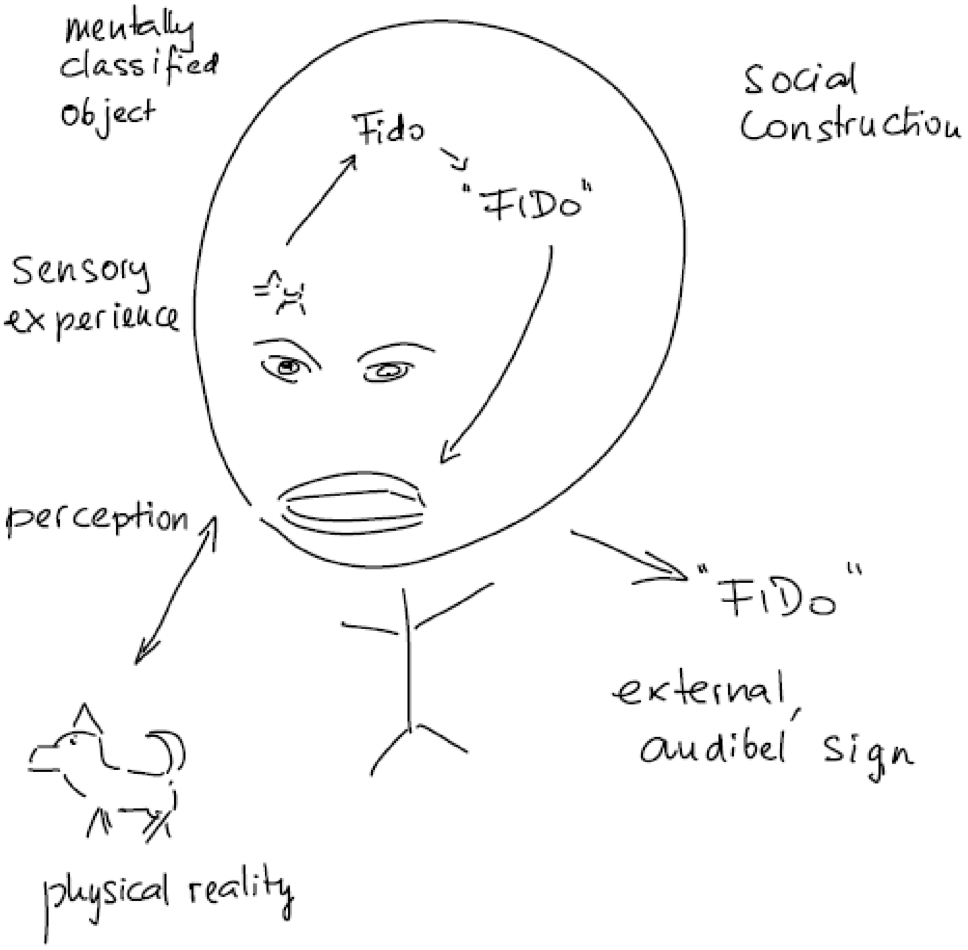

Figure 6: The grounding of constructs in experiential concepts

The representable signs are constructed as models of reality. These signs may be verbal descriptions, oral or written, computational models, sketches, etc. They are strongly interconnected by operations and relations. I describe such models as algebras and posit that they are—in a fuzzy way—homomorphic to reality (Lawvere et al. 2005; Kuhn 2007).

The "fuzzy homomorphism" between experience and mental models which must be reflected in the verbal communication seems to be sufficient to converge into a common encoding over repeated experiences. The fact that initial language acquisition occurs in a simplified reality and within a supportive affective environment may significantly influence how the mechanism of language acquisition works.

3.3.1 Context

The meaning of constructions are determined in a web of concepts that are bound by the relations between the constructs. The full set of concepts that are interrelated are called the context of the construct; the semantics of the construct is determined only through the relations in this context and within this context. Notice the terminology: a person is in a real world *situation*, the meaning of a sign (construct) is given by *context*.

Considering these structures as algebraic structures indicates that the semantics is determined only up to a structure preserving isomorphism. This is not a limitation and an uncertainty but is the precondition for communication to be possible: it must be possible to translate between different representations (mental, verbal, written). To maintain the meaning; the translations must be structure preserving mappings (Eco 2003).

3.3.2 Grounding

Using Searle’s formula for a semi-formal treatment, I posit that mental experiential concepts have corresponding representable concepts; the formal “X counts as Y in context Z” is generalized: an experiential concept X counts as a representable concept Y. Note that the experiential concept—an experience of a thing in reality—can be caused by an ordinary physical object or a physical object that is intended as a sign (Eco 1976). The formula provides grounding for all constructions in mental concepts, which are all directly or indirectly related to such grounding items and through these experientially grounded. Ungrounded (“freestanding Y terms” (Zaibert et al. 2004)) would be meaningless (in the sense of Wittgenstein (1960)).

3.3.3 Communication

Despite the fact that we do not know exactly how humans learn their mother tongue (Eco 1976; Pinker 1995) it is an empirical observation that humans establish a consensus on the meaning of external signs. Human communication is possible, even though it is not perfect! Acquiring a language means to establish a correspondence between experiential concepts and constructions.

3.3.4 Imperfections in communication

The meaning of a sign is defined in its context and this context can vary between sender and recipient of a sign. If a sign is unique to a context, no confusion within a context is possible, but for homonyms (same sign but different meaning) and polysemy, where the same word means in different contexts different things and a potential for imperfect communication exists. WorldNet (Fellbaum 1998) documents polysemy in natural language by separating different meanings of a word in synsets.

The imperfection of communication increases with the distance between the contexts. A description of a soil for civil engineering, hydrology, or agriculture may use the same words, but the meaning is different because the words are in each science connected differently; the structure established in each context is different. In the hydrologic or agricultural context international organizations, e.g., ISO, establish formalized contexts in which signs can be exchanged with the assumption of unchanged semantics.

In normal communication circumstances, multiple contexts are combined. For example, participants in a meeting each have a subjective component in their context as well as a role influenced context; much discussion in meetings serves to align the contexts of the participants (Rottenbacher 2006).

Language classifies (constructions of) objects in *is_a* taxonomies; this classification is very similar and reflects similar concerns as mental classification (subsection 2.3.3). Separating mental and linguistic classification is a step to clarify the ontology of information process.

The theory of supervaluation gives guidelines how to deal with the integration of multiple context and reasoning in an integrated data collection. If the semantics of the context are available as formally described ontologies (e.g., in OWL (Dieckmann 2003)) then formal conversions between codes from different contexts can be attempted, but general solutions are not yet known. The translation between contexts must always relate the signs in the context back to the grounding items and then forward to the sign of interest in the other context (Figure 7).

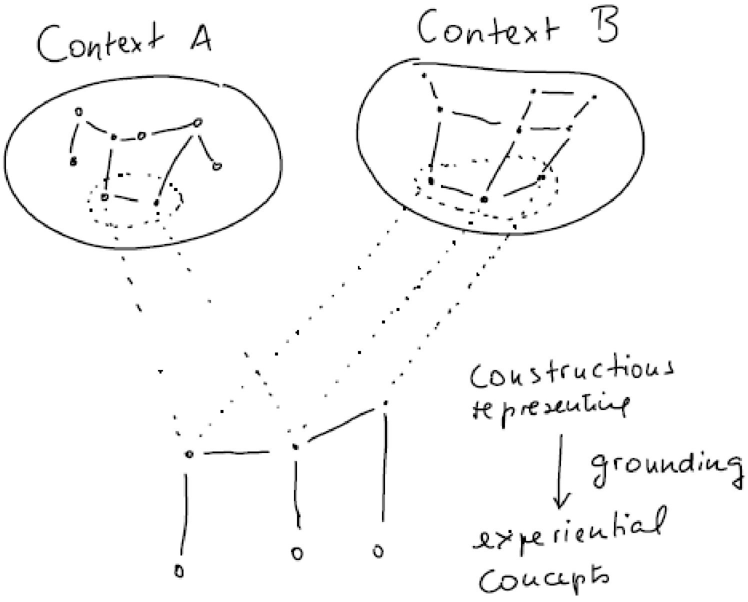

Figure 7: Translating between contexts in one individual through experientially grounded concepts

3.3.5 Constructions are without imperfections

Constructions are, unlike observations (in tier1) descriptions or classification of objects (in tier2) without imperfection and error, as long as they are used in a fixed and shared context. As an everyday example, consider a description of the paper bill in (Figure 8):

Figure 8: Example of socially constructed objects

The length of the paper is a physical (tier 2) observation: 134 mm, with a standard deviation of 3/100 mm. Once established that this is constructed as a Czech bill of 50 Czech Crowns; there is no uncertainty, the value in the Czech context is not 49.90 or 50.05 with any probability! If we leave the Czech context, then the value expressed in Euro may be uncertain, today the exchange rate is 33.2050 Crowns per Euro, which gives an approximate value of 1.50 Euro for the bill.

The value 50 is here—unlike the measured physical length—without error, directly contradicting the often heard statement that “all data contains some error”. It is correct only for tier 1 and 2, but not for the constructions of tier 3. The imperfections of tier 3 are introduced by

- establishing the connection between experiential concepts and constructions—the ‘subsumption’ of the law where one establishes whether a concrete act was ‘murder’ or ‘manslaughter’, and
- translating between contexts.

4 Use of Imperfect GIS Data for Decisions

A GIS is built to have data available preparing for future decisions the data is imperfect and the decision will be affected by these imperfections. The rules of error propagation, fuzzy set, Bayesian networks, and logical inferences to align the semantics have been developed to reduce the effects of imperfections on the decision (Morgan et al. 1998). These computational approaches are complex and require large amount of data. They are based on the assumption that with complete and perfect information our decisions would be perfect and a minor related assumption that more work will lead to better decisions (Gigerenzer et al. 1999). Surprisingly, this assumption is not always true—complex models tend to “overfit” and use elements present in the past to predict the future incorrectly. Equally important, real world decision making has always only limited resources of data and time available. Gigerenzer and his group have shown that despite such limitations, decisions are often correct—and leaving opportunities for later adaptation is wise in an ever changing, uncertain world (Popper 1984). Gigerenzer and his group argue that human behavior is adapted to the world and exploits the structure of the world—physical and social—to make good decisions. Efforts that blindly try to improve the data independent of how it is used to make decisions related to the highly structured world, are misguided and a simple loss of resources.

In this section I show a few compensatory effects that are built into data collection and decision strategies that help to explain why simple strategies using imperfect data are still effective.

4.1 Correlation

The aspects of the world relevant for human life are highly spatially and temporally autocorrelated. The first law of geography by Waldo Tobler says: "Everything is related to everything else, but near things are more related than distant things." (Tobler 1970, 236). This has several compensatory effects:

- Measurement errors can be reduced by filling.
- Desirable data that is not available can be replaced by correlated data.
- Only data of nearby objects and of similar spatial and temporal frequency as the decision is affecting is relevant. Processes that are much faster or much slower, or which are much larger or much smaller in space are constant influences and not included as time varying values (i.e., processes).

Factors that influence a decision in a minor way can be neglected; especially factors the influence of which is small compared to the imperfections of other factors.

4.2 Granulation and Classification

It is well known that classification of objects by humans is a complex and multifaceted process; here I only address the cognitive (subjective, personal) mental classification of objects with respect to a potential interaction not the taxonomy fixed in a vocabulary.

Empirical evidence shows that mirror neurons (Rizzolati et al. 2002) found in humans and (at least) apes classify not only operations the cognizant agent sees (i.e., visually perceives) but also classify the objects with respect to having the right properties to be involved in an operation. Potential interactions between the agent and objects or interactions of interest between objects produces conditions these objects must fulfill, expressed as a property and a range for the value of the property. In this way operations of interest indicate what properties are important and these important properties are then used to determine boundaries of objects (above subsection 3.2.3) (Frank 2006; Kuhn 2007).

4.3 Constructions in Legal Decisions

Legal or administrative decisions do not tolerate errors; they are typically subject to review by others; how are endless arguments due to limitations in the data avoided?

In the context of the legal or administrative decision, the data are well defined and without imperfections; the imperfections of the observations, granulation and classification are reduced by procedural description how the relevant properties must be observed and classified, often concluded with an authoritative statement about the subsumption of a real world fact under a construction(the rules of “due process” (Black 1996).

4.4 Communication

Previous research in the semantics of linguistic classification has identified a radial structure (radial categories (Rosch 1973)). The same in multiple contexts has widely overlapping applicabilities, which share a core meaning (Figure 9).

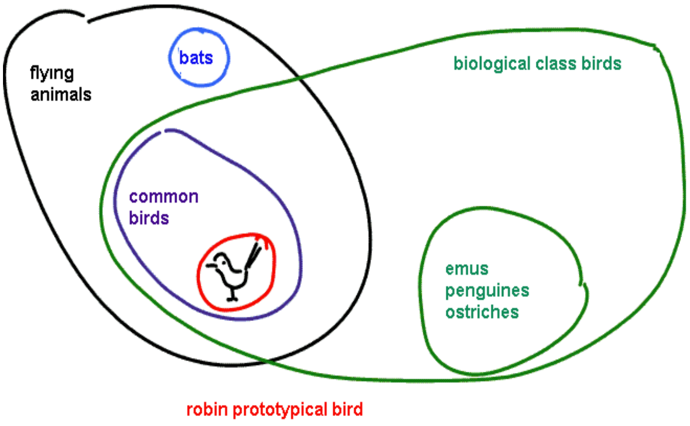

Figure 9: Different meaning of ‘Bird’ in different contexts

Exemplars in the area of overlap are coded in each context with the same code. Other exemplars are coded differently, depending on the context. One speaks of “better examples” for a class (e.g., robin of sparrow are “better” birds than penguin or ostrich), which is in contrast to a set theoretic approach, where an exemplar is or is not member of the class and no gradual membership is possible.

The rich structure of reality as we experience it is approximately shared by all humans (Lakoff 1987), because they share a large part of daily experiences (eating, drinking, sleeping, etc.). This approximation of the experiential grounding is usually sufficient to establish a mapping between structure encountered in a text received and our own structure among constructs.

4.5 A Linguistic Classification

Most classifications are done when a noun is applied to an object; we talk about buildings, dogs, and trees, implicitly classifying the objects we see as belonging to the class of Buildings, Dogs, or Trees. Such classifications are of a different type than the strict or fuzzy classification because the classes used in linguistic classification do not divide the world in jointly exhaustive and pair wise disjoint classes (or the relaxed form of fuzzy classification, where membership sets overlap but membership degree add up to 1). Linguistic classes show prototype effects (Rosch 1978).

Concept maps show clearly the relations between them, at least for a single view point. In a context some aspects may dominate and others are less important. Last, but not least, boundaries between natural language concepts (not legal or scientific constructs) tend to be at natural breaks: the distinction between cats and dogs is unproblematic, because no intermediate individuals exist (horses and donkeys provide one of the rare counterexamples!).

5 Active Methods to Compensate for Imperfections

The data we use to make decisions include imperfections, but we use it nevertheless and the decisions are usually better because we had the information than not. The question is therefore what are methods used to reduce the negative effects.

5.1 Multiple Observations

Surveyors observe more and more often than absolutely necessary to obtain a desired value. The redundancy created by repeated measures is used to, first, reduce the (statistical) imperfection in the value; second to assess the level of statistical imperfections (i.e., the standard derivation of the PDF) and, third, to check against blunders and therewith to improve the reliability of the value.

5.2 Safety Margins for Decisions

If you have 2 liter of gas in the tank of your car, which uses 10 l/100km, you will not pass a filling station when you know that the next is 17 km later. We should be able to reach the next station, but the safety margin seems too low to risk to remain with no gas on the road. Similarly, most engineering decisions include safety margins—in our example a mental rule that says for example 'you should always have 5 l gas in your car'. Safety margins in engineering decisions are factors that increase the load and decreases the resistance of the materials assumed, and can be translated to a probability of failure (often 1 or 2% failure probability is acceptable) (Schneider 1995; Frank to appear 2008).

5.3 Absorption

Life is risky and many events cannot be predicted with accuracy at reasonable cost, but it is sometime possible to find somebody else to shoulder the risk for us. Bédard has called this 'uncertainty absorption' (Bédard 1986). It comes in many forms:

Insurance: I pay another party to cover the cost of errors in my decision (e.g., fire or flooding of my home).

Guarantee: Another party guarantees the subsumption that led to a construction. Certain data in some registers, for example, the German or Swiss cadastre (Grundbuch) is by law correct and the state guarantees it.

Liability: Another party is paid to make a subsumption for me and liable for the risks involved. Many types of certificates of professionals, e.g., building inspections certificates, include liability of the professional making the judgment. The professional often carries insurance to cover the rare events of error.

6 Conclusion

A systematic study of how imperfections are introduced in the data can be used to identify three large groups of types of imperfections

- *probability* related to observations of physical reality,
- *fuzzy set* descriptions of classification and subsumption, and
- *supervaluation* methods to cope with change of context of constructions.

Considering that humans have adapted their methods to make decisions to their environment, what Gigerenzer calls ecological rationality, even with the imperfect data in a GIS useful decisions can be made. Several methods are used to reduce the negative effects of imperfections on decisions, some based on properties of physical reality (e.g., correlation) and others socially engineered (e.g., constructions).

References

Abler, R. (1987). *Review of the Federal Research Agenda*. International Geographic Information Systems (IGIS) Symposium (IGIS'87): The Research Agenda, Arlington, VA.

Black, H. C. (1996). *Black's Law Dictionary,* West Publishing.

Burrough, P. A. (1996). Natural Objects with Indeterminate Boundaries. *Geographic Objects with Indeterminate Boundaries*. P. A. Burrough and A. U. Frank. London, Taylor and Francis**:** 3-28.

Couclelis, H. (1992). People Manipulate Objects (but Cultivate Fields): Beyond the Raster-Vector Debate in GIS. *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. A. U. Frank, I. Campari and U. Formentini. Berlin, Springer-Verlag. **639:** 65-77.

Dieckmann, J. (2003). DAML+OIL und OWL XML-Sprachen für Ontologien. Berlin**:** 21.

Eco, U. (1976). *A Theory of Semiotics*. Bloomington, Indiana University Press.

Egenhofer, M. and A. U. Frank (1986). *Connection between Local and Regional: Additional "Intelligence" Needed*. FIG XVIII International Congress of Surveyors, Toronto, Canada (June 1-11, 1986).

Fellbaum, C., Ed. (1998). *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. Cambridge, Mass., The MIT Press.

Frank, A. U. (2001). "Tiers of Ontology and Consistency Constraints in Geographic Information Systems." *International Journal of Geographical Information Science* **75**(5 (Special Issue on Ontology of Geographic Information)): 667-678.

Frank, A. U. (2003). Ontology for Spatio-Temporal Databases. *Spatiotemporal Databases: The Chorochronos Approach*. M. Koubarakis, T. Sellis and e. al. Berlin, Springer-Verlag. **2520:** 9-78.

Frank, A. U. (2006). *Distinctions Produce a Taxonomic Lattice: Are These the Units of Mentalese?* International Conference on Formal Ontology in Information Systems (FOIS), Baltimore, Maryland, IOS Press.

Frank, A. U. (to appear 2008). "Analysis of Dependence of Decision Quality on Data Quality." *Journal of Geographical Systems*: 18.

Ganter, B., G. Stumme and R. Wille, Eds. (2005). *Formal Concept Analysis Foundations and Applications*. Berlin, Heidelberg, Springer.

Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, NJ, Lawrence Erlbaum.

Gigerenzer, G., P. M. Todd and e. al. (1999). *Simple Heuristics That Make Us Smart*. New York, Oxford University Press.

Goodchild, M. F., M. J. Egenhofer, K. K. Kemp, D. M. Mark and E. Sheppard (1999). "Introduction to the Varenius Project." *International Journal of Geographical Information Science* **13**(8): 731-745.

Guarino, N. and C. Welty (2000). Ontological Analysis of Taxonomic Relationships. *Proceedings of ER-2000, 19th Int. Conference on Conceptual Modeling*. E. Laender and V. Storey, Springer-Verlag.

Heidegger, M. (1927; reprint 1993). *Sein und Zeit*. Tübingen, Niemeyer.

Husserl (1900/01). *Logische Untersuchungen*. Halle, M. Niemeyer.

Kuhn, W. (2007). *An Image-Schematic Account of Spatial Categories*. 8th International Conference, COSIT 2007, Melbourne, Australia, Springer.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL, University of Chicago Press.

Lawvere, F. W. and S. H. Schanuel (2005). *Conceptual Mathematics: A First Introduction to Categories*. Cambridge, Cambridge University Press.

McCarthy, J. and P. J. Hayes (1969). Some Philosophical Problems from the Standpoint of Artificial Intelligence. *Machine Intelligence 4*. B. Meltzer and D. Michie. Edinburgh, Edinburgh University Press**:** 463-502.

Morgan, M. G. and M. Henrion (1998). *Uncertainty A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press.

NCGIA (1989a). "The U.S. National Center for Geographic Information and Analysis: An Overview of the Agenda for Research and Education." *IJGIS* **2**(3): 117-136.

NCGIA (1989b). Use and Value of Geographic Information Initiative Four Specialist Meeting, Report and Proceedings, National Center for Geographic Information and Analysis; Department of Surveying Engineering, University of Maine; Department of Geography, SUNY at Buffalo.

Openshaw, S., M. Charlton and S. Carver (1991). Error Propagation: A Monte Carlo Simulation. *Handling Geographical Information*. I. Masser and M. Blakemore. Essex, Longman Scientific & Technical. **1:** 78-101.

Pinker, S. (1995). *The Language Instinct*. New York, HarperPerennial.

Popper, K. R. (1984). *Logik der Forschung*. Tübingen, J. C. B. Mohr.

Randow, G. v. (1992). *Das Ziegenproblem - Denken in Wahrscheinlichkeiten*. Hamburg, Rowohlt Taschenbuchverlag.

Raubal, M. (2002). *Wayfinding in Built Environments: The Case of Airports*. Münster, Solingen, Institut für Geoinformatik, Institut für Geoinformation.

Rizzolati, G., L. Craighero and L. Fadiga (2002). The Mirror System in Humans. *Mirror Neurons and the Evolution of Brain and Language*. M. Stamenov and V. Gallese, John Benjamins Publishing Company**:** 37-59.

Rosch, E. (1973). "Natural Categories." *Cognitive Psychology* **4**: 328-350.

Rosch, E. (1978). Principles of Categorization. *Cognition and Categorization*. E. Rosch and B. B. Lloyd. Hillsdale, NJ, Erlbaum.

Rottenbacher, C. (2006). *Bewegter Planungsprozess*. Vienna, Technical University Vienna. PhD.

Sartre, J. P. (1943; translated reprint 1993). *Being And Nothingness*. New York, Washington Square Press.

Schneider, M. (1995). *Spatial Data Types for Database Systems*. Hagen, FernUniversitaet.

Searle, J. R., Ed. (1995). *The Construction of Social Reality*. New York, The Free Press.

Simon, H. (1956). "Rational Choice and the Structure of the Environment." *Psychological Review* **63**: 129-138.

Smith, B. and P. Grenon (2004). "SNAP and SPAN: Towards Dynamic Spatial Ontology." *Spatial Cognition and Computing*(4): 69-103.

Tobler, W. R. (1970). "A Computer Model Simulation of Urban Growth in the Detroit Region." *Economic Geography* **46**(2): 234-240.

Tomlin, C. D. (1983). *A Map Algebra*. Harvard Computer Graphics Conference, Cambridge, Mass.

Wittgenstein, L. (1960). *Tractatus logico-philosophicus*. London, Routledge & Kegan Paul.

Zadeh, L. A. (1974). "Fuzzy Logic and Its Application to Approximate Reasoning." *Information Processing*.

Zadeh, L. A. (2002). Some Reflections on Information Granulation and Its Centrality in Granular Computing, Computing with Words, the Computational Theory of Perceptions and Precisiated Natural Language. *Data Mining, Rough Sets and Granular Computing*. T. Y. Lin, Y. Y. Yao and L. A. Zadeh. Heidelberg, Germany, Physica-Verlag GmbH**:** 3-20.

Zaibert, L. and B. Smith (2004). Real Estate - Foundations of the Ontology of Property. *The Ontology and Modelling of Real Estate Transactions: European Jurisdictions*. H. Stuckenschmidt, E. Stubkjaer and C. Schlieder, Ashgate Pub Ltd**:** 35-51.