

Experiences in Developing Landform Ontologies

R. K. Straumann

Department of Geography, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
Telephone: +41 44 635 51 98
Fax: +41 44 635 68 48
Email: ralph.straumann@geo.uzh.ch

1. Landforms – Ontologies and Extraction

The aim of this paper is to briefly review some of the work in landform-related research, to highlight the need for a sound ontological basis to such efforts and to present and discuss an approach for gathering domain knowledge and the problems encountered.

There is a considerable volume of geomorphometric literature centred on topographic eminences, specifically publications regarding the delineation or extraction of hills or mountains; e.g. Fisher et al. (2004), Chaudhry and Mackaness (2007) and (less spatially) Greatbatch et al. (2007). However, many of the approaches to landform delineation/extraction start to rapidly delineate crisply or extract fuzzily the desired objects. The authors rely to a certain degree on some common knowledge about what a mountain, a hill or a range is and, to a certain degree, assume that these concepts match what they extract. But there is fundamental work which acknowledges that these and similar landform concepts are not clear at all and that research into concepts and their formalisation is needed (e.g. Brändli 1996, Schmidt and Dikau 1999). A strain of research tries on a fundamental level to elucidate the ontology of geographic objects and with these, of landforms (e.g. Smith and Mark (2001, 2003), Smith and Varzi 2000, Mark and Smith 2004, Mark and Sinha 2006). Furthermore Mark et al. (2007) set out clearly why conceptualisations of landforms are not alike for people of different cultures or language groups.

In Geographic Information Science and geomorphometry there are a range of researchers who have developed methods to describe earth surface forms from digital elevation models (DEMs). Some of this interest has turned away from the mere description of predefined areas of land through DEM derivatives such as hypsometry, gradient or aspect to semantically richer characterisations of surface form. There is a huge breadth of publications related to the extraction of units that are homogeneous in relation to some surface properties. These units are termed (among others) “landform elements”. Besides these landform elements, there has been growing interest in recent years concerning landforms, i.e. larger regions of similar form character.

Besides, geographic information science has seen some effort to render geographic information systems more usable for lay-persons (e.g. Mennis et al. 2000). This has spawned a considerable amount of work on fuzzy spatial relations, such as something “being near” something else (e.g. Robinson 2000). Given this background, we think a similar case can be made that in the long run it would be valuable to enable GISs to make sense of landform terms such as “valley” or “mountain”.

We thus argue that – although there are already many extraction and classification algorithms especially for what are termed landform elements – there is a need to further strengthen the ontological basis such approaches implicitly rely upon. It might

be insightful to review the breadth of landforms, to characterise their properties and interrelationships before pondering about ways to extract them from DEMs. We feel such an endeavour may serve well in order to strengthen efforts towards their automatic extraction, specifically, and towards landscape characterisation from DEMs in general.

2. Candidate Sources of Landform Description Catalogues

To elucidate the landform-related domain knowledge of geomorphology we drew on a number of sources. The reference works used range from standards dealing with (among others) landforms to more ontological sources. To build a collection of landform terms that represent important concepts in the domain knowledge, the listed data sources need to be scanned for landform-related categories.

WordNet (2006) is a lexical database of the English language held at the Princeton University. Word types are grouped into synsets. Synsets represent cognitive synonyms and each synset stands for a distinct concept. Different synsets are interlinked by various relations such as hypernymy (superordination), hyponymy (subordination), holonymy (whole-to-part relation) and meronymy (part-to-whole relation). In WordNet most landforms seem to be hyponyms of the synset „geological formation, formation”.

SDTS (Spatial Data Transfer Standard; USGS 2007) has been devised as a means of transferring spatial data between computer systems. It is made up of base specifications and profiles. SDTS was ratified by the American National Standards Institute (ANSI) in 1998. We investigated SDTS mainly along the lines of Mark and Smith (2004: 82) who singled out 26 landform categories appearing “to fall under the broad superordinate category of ‘landform’”.

DIGEST (Digital Geographic Information Exchange Standard; DGIWG s.a.). The Digital Geospatial Information Working Group (DGIWG) was established in 1983 in order to develop standards for the exchange of geographical information among NATO members. DIGEST as developed by DGIWG has become a NATO Standardization Agreement. Landform categories are primarily contained in the section “D-Physiography – DB-Physiography-Landforms” of the Feature and Attribute Coding Catalogue.

Alexandria Digital Library (ADL) Feature Type Thesaurus (Alexandria Digital Library Project 2004). The ADL Feature Type Thesaurus has been developed for typing entries in the ADL Gazetteer and as a means to foster gazetteer interoperability. It contains a hierarchical listing of terms in the administrative, hydrographic, land parcels, man-made, physiographic and regional places domains. Landform-related categories can be found in the domains “Physiographic features” and “Hydrographic features”.

SUMO (Suggested Upper Merged Ontology; SUMO 2009, Niles and Pease 2001) is an upper (or top-level, foundation) ontology, i.e. an ontology of very general concepts that are shared among all domains. It is a candidate ontology for the Standard Upper Ontology (IEEE SUO Working Group 2003). The whole SUMO consists of the core SUMO itself, the Mid-Level Ontology (MILO) and several domain ontologies among which there is also one for geography. For our investigation we exploited all concepts that are subclasses of the concept “LandForm” in this geography ontology.

3. Problems in Building a Taxonomy/Ontology

In distilling domain knowledge about landforms and a hierarchy of landforms from the afore-mentioned works and supported by geomorphological literature we found several difficulties. Among others, obvious difficulties arose with ambiguities in definitions or contradictions between definitions. There are also certain categories which suffer from under-specification. In the remainder of this paper we will detail some of these issues.

3.1. Ambiguities Regarding Form and Levels of Granularity

In most reference works the coverage of (sand) dunes is relatively shallow, i.e. they often contain only reference to a category “dune” or “sand dune”. This leaves much room for ambiguity regarding the form of such features. However, for the “dune” category this seems inherent since it is (at least at that general level) mainly characterised by material (sand or granular material) and process (wind-blown), but not form. Both the reference works and other geomorphologic literature heavily stress these aspects. This situation is in some respect similar to that of moraines which are also predominantly defined by material (boulders, stones, debris) and process (transport and deposition by a glacier) but show a large variety of forms.

However, a definition for dunes as “ridges or hills of sand” (DIGEST) is of limited use for devising a delineation/extraction method since it allows for much ambiguity regarding form. In such cases we suggest the enrichment of the vocabulary provided by the reference works to obtain a representation at a sufficient granularity – e.g. through inclusion of categories such as “transverse dune”, “longitudinal dune”, “barchan dune” etc. which are more specific regarding form (and context). These can be grouped in somewhat artificial categories such as “ridge-shaped dunes” or “hill-shaped dunes”.

3.2 Under-specification

Under-specification in the sense of lacking indications as to e.g. the typical size of instances of some category is common. However, in what follows we want to discuss an under-specification which while failing to account for a threshold size or differentiating property at least provides an ordinal measure for distinguishing two categories. While this situation is better than general under-specification, it still has adverse implications for work on delimitation/extraction of such features.

The two probably most popular examples of topographic eminences are “mountain” and “hill”. However, there is an obvious conceptual uncertainty, since the dichotomy between the two is unclear or underdeveloped. In reference works mountains are often described as being “higher than a hill”, whereas hills are described as being “smaller” or “shorter” than mountains (e.g. WordNet, DIGEST, SUMO). However, the uncertainty regarding their semantic delimitation does not seem to make people feel uncomfortable using the two terms. On contrary, the two terms are very popular with non-experts (cf. e.g. Battig and Montague 1969 as cited in Smith and Mark 2001, Smith and Mark 2001).

The conceptual uncertainty regarding hills and mountains is probably due to the (relative to “hill”) late introduction of “mountain” into the English language. According to the Oxford English Dictionary (OED; Oxford University Press s.a.), “hill” was formerly the all-encompassing term “including what are now called mountains”. After the introduction of “mountain”, however, “hill” was “gradually restricted to heights of less elevation”. During the 18th century “mountain” was still used to designate objects of moderate altitude (OED lists a quotation referring to St. Germain near Paris being situated on a mountain).

Derungs and Purves (2007) empirically investigated the conception of mountains of Swiss citizens. Terms often associated with mountains were e.g. “high”, “rock”, “snow”, “steep”. The question regarding minimum altitude of a feature to be called a mountain resulted in a very broad distribution of the answers (1364(±713)metres) hinting indeed that altitude is not a useful criterion.

The solution to this problem of delimitation could be for geomorphometric delineation/extraction tasks to resort to a superordinate category (e.g. “topographic eminence”) and leave the differentiation between hill and mountain to the user (e.g. via a user-adaptable thresholding process) or for later when more research into the differentiation (e.g. using other attributes such as ruggedness or landcover) has been done.

4. Reflections

This paper has touched upon some of the issues encountered with gathering and reconciling geomorphologic knowledge. We are convinced that making such knowledge more explicit before devising geomorphometric methods to extract features of interest could improve present approaches to surface form characterisation. Such knowledge can be used to probably first extract ‘cores’ of landforms (e.g. valley floors for valleys (Straumann and Purves 2008) or peaks and ridges for mountains (Mark and Sinha 2006)). Adopting such an approach subsequently needs methods to be found in order to sensibly ‘spread’ conceptual cores for finding the extent of a landform.

References

- Alexandria Digital Library Project, 2004, ADL Gazetteer. <http://www.alexandria.ucsb.edu/gazetteer>.
- Battig WF and Montague WE, 1969, Category norms for verbal items in 56 categories: a replication and extension of the Connecticut Norms. *Journal of Experimental Psychology Monograph*, 80(3), Part 2: 1-46.
- Brändli M, 1996, Hierarchical models for the definition and extraction of terrain features. In: Burrough PA and Frank AU (eds), *Geographic Objects With Indeterminate Boundaries, GISDATA II*, Taylor & Francis, London, UK, 257-270.
- Chaudhry O and Mackaness W, 2007, Automatic identification of hills and ranges using morphometric analysis. *Proceedings of GISRUK 2007*, Maynooth, Ireland, 333-342.
- Derungs C and Purves RS, 2007, Empirical experiments on the nature of Swiss mountains. In: Winstanley AC (ed), *Proceedings of GISRUK 2007*, Maynooth, Ireland, 348-352.
- DGIWG, s.a., Digital Geographic Information Exchange Standard (DIGEST). <http://www.dgiwg.org/digest>.
- Fisher P, Wood J and Cheng T, 2004, Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. *Transactions of the Institute of British Geographers*, 29(1): 106-128.
- Greatbatch I, Wood J and Fisher P, 2007, A comparison of morphometric and web prominence of mountain features. *Proceedings of GISRUK 2007*, Maynooth, Ireland, 312-317.
- IEEE SUO Working Group, 2003, Standard Upper Ontology Working Group (SUO WG) Home Page. <http://suo.ieee.org/index.html>.
- Mark D and Sinha G, 2006, Ontology of landforms: Delimitation and classification of topographic eminences. In: Raubal M, Miller HJ, Frank AU and Goodchild MF (eds), *Proceedings of GIScience 2006*, Muenster, Germany, 129-132.
- Mark DM and Smith B, 2004, A science of topography: From qualitative ontology to digital representations. In: Bishop MP and Shroder Jr JF (eds), *Geographic Information Science and Mountain Geomorphology*, Springer, Berlin, Germany, 75-100.
- Mark DM, Turk AG and Stea D, 2007, Progress on Yindjibarndi Ethnophysiography. In: Winter S, Duckham M, Kulik L and Kuipers B (eds), *Proceedings of COSIT 2007*, LNCS, 4736, 1-19.
- Mennis JL, Peuquet DJ and Qian L, 2000, A conceptual framework for incorporating cognitive principles into geographical database representation. *International Journal of Geographical Information Science*, 14(6): 501-520.
- Niles I and Pease A, 2001, Towards a Standard Upper Ontology. In: Welty C and Smith B (eds), *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, FOIS 2001*, Ogunquit, USA, 2-9.

- Oxford University Press, s.a., Oxford English Dictionary (OED) Online. <http://dictionary.oed.com>.
- Robinson VB, 2000, Individual and multipersonal fuzzy spatial relations acquired using human-machine interaction. *Fuzzy Sets and Systems*, 113(1): 133-145.
- Schmidt J and Dikau R, 1999, Extracting geomorphometric attributes and objects from digital elevation models – semantics, methods, future needs. In: Dikau R and Saurer H (eds), *GIS For Earth Surface Systems*. Gebrüder Borntraeger, Berlin, Germany, 153-173.
- Smith B and Mark DM, 2001, Geographical categories: An ontological investigation. *International Journal of Geographical Information Science*, 15(7): 591-612.
- Smith B and Mark DM, 2003, Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design*, 30(3): 411-427.
- Smith B and Varzi AC, 2000, Fiat and bona fide boundaries. *Philosophy and Phenomenological Research*, 60(2): 401-420.
- Straumann RK and Purves RS, 2008, Delineation of valleys and valley floors. In: Cova TJ, Miller HJ, Beard K, Frank AU and Goodchild MF (eds), *Proceedings of GIScience 2008*, LNCS, 5266, Park City, USA, 320-336.
- SUMO, 2009, The Suggested Upper Merged Ontology (SUMO), Ontology Portal. <http://www.ontologyportal.org>.
- USGS, 2007, Spatial Data Transfer Standard Home Page. <http://mcmcweb.er.usgs.gov/sdts>.
- WordNet, 2006, WordNet 2.1 for Windows. A lexical database for the English language. <http://wordnet.princeton.edu>.