From Land Form to Landforms: Bridging the Quantitative-Qualitative Gap in a Multilingual Context

D. M. Mark[1]

[1]NCGIA & Department of Geography, University at Buffalo, Buffalo, NY 14261 USA
Telephone: (1) 716 645 0505
Fax: (1) 716 645 5957
Email: dmark@buffalo.edu

1. Introduction

Since the 1960s, geomorphometry has emphasized geometric calculations based on local operators applied to digital elevation data (Evans, 1972; Mark, 1975). Digital elevation data represent discrete approximations to fields of elevations that represent the objective, measurable form of land surfaces.

With the emergence of the World Wide Web and the Semantic Web, there has been increased need for methods to conversion of elevation fields into cognitively meaningful landforms. Some of the research has approached the problem in a 'bottom-up' way, classifying form at a local level. However, in a somewhat counter-intuitive development, improving quality and resolution of digital elevation has had the effect of broadening the conceptual and computational gap between local geometry and meaningful landforms. The issue of automated feature extraction and classification is further complicated by the fact that people from different cultures, speaking different landscapes, group landforms into categories in different ways. Procedures for the automated detection, delimitation, and classification of landforms from elevation data may themselves need to be different for different languages, and at the very least will need to have different parameters.

This paper expands on the issues outlined above, and then present evidence for cultural and linguistic differences in the classification of landforms. Then some aspects of the solutions will be presented.

2. Ontology and Feature Extraction

2.1 Ontology of Landforms

A comprehensive ontology of landforms appears to consist of several distinct components. One part would be an ontology of fields (Peuquet et al., 1998). Mathematically, the shape of the Earth's solid surface can be approximated by a single-valued field (Mark and Smith, 2004). Then, various discrete representations of altitude, such as elevation matrices and TINs, can be considered to be discrete approximations to this field.

Next, we have the question on the ontological nature of landforms themselves. Strictly speaking, landforms are parts of the Earth's surface. But not arbitrary parts. To be a landform, a part of the Earth's surface must have some coherence of form (shape) or process or both (Smith and Mark, 2003). In the DOLCE ontology, landforms are probably best considered to be *features* of a planetary surface. "Typical examples of features are 'parasitic entities' such as holes, boundaries, surfaces, or stains, which are generically constantly dependent on physical objects (their hosts)" (Masolo et al., 2003, p. 29). They go on to mention that "some features may be relevant parts of their host, like a bump or an edge, or places like a hole in a piece of cheese … which are not parts of their host." Development of this ontology will provide elements of the Upper Level or Foundational Ontology for the geographic domain.

A third component of ontology of landforms is the taxonomy of types or kinds. What kinds of landforms are recognized, and how do the kinds relate to each other? For one thing, do landforms constitute a single domain of reality, or of the perceived human environment? If so, what are the main subdivisions? Convex and concave come to mind, but many parts of the Earth's surface are approximately planar. Do horizontal planar regions constitute a top-level class of landforms? Are vertical planar regions (cliffs in English) forms in their own rights, or parts of concave or convex landforms, or both? Voegelin and Voegelin (1957, p. 13) recognized "the Domain of Topography" as a top-level domain, and divided topography into three sub domains: eminences in the landscape; longitudinal depressions in the landscape; and oval or round openings in the earth. Eminences are regions that are higher than their surroundings. (A convex landform could protrude from the side of a hill while not being an eminence, but large examples of this are rare.) The main types of eminences in English are mountains and hills, but the English language also includes terms such as ridge, range, plateau, mesa, butte, pinnacle, tor, and others, which also denote types of eminences. Research into semantic similarity is needed to determine whether these basic-level categories are arranged in a hierarchical fashion, or whether all of these are simply types of eminence. A similar situation exists for longitudinal depressions in the landscape.

Lastly, the ontology of fields must be integrated with the taxonomy of landform types, all within an Upper Level ontology such as DOLCE—not an easy task!

2.2 Cultural Differences

Given the ontological situation outlined above, it is clear that there is room for different human cultures and speech communities to 'parse' the same landscape into different features, and also to group those features into named categories in different ways. Recently, Mark, Turk, and colleagues have presented empirical evidence of this, based on ethnographic studies with the Yindjibarndi and Navajo people (Mark and Turk, 2003; Mark et al., 2007). One does not, however, need to study Indigenous languages to find examples. For example, Mark (1993) presented an example of different categories for water bodies in English and French.

In the Yindjibarndi language, spoken mainly around Roebourne, Western Australia, there is a single term, ***marnda***, that refers to entities that in English would be called rock (the material), hills, mountains, ridges, ranges, etc. (Turk and Mark, 2008). ***Marnda*** appears to be used to refer to almost any eminence. Rather than having separate terms for large or small or flat-topped eminences, Yindjibarndi speakers combine the term ***marnda*** with general adjectives that denote size or shape of the feature. The Yindjibarndi language does have a few other terms that refer to convex features of the landscape (Turk and Mark, 2008). ***Burbaa*** apparently refers to a relatively small ***marnda*** that has a smooth, rounded shape. ***Bargu*** can refer to a small ***marnda*** with a rough or irregular profile, or to a rough rock outcrop protruding from the side of a hill (not an eminence). Lastly, a ***bantha*** is material piled up by a person or animal, and ***munggu*** refers to a termite mound. The distinctions between these categories do not line up well with distinctions among types of eminences made by English speakers.

Preliminary results for the Navajo language, not yet published, indicate that Navajo speakers have terms for eminences that differ in meaning from both the Yindjibarndi terms and the English terms (Mark, Stea, Topaha, and Turk, unpublished). For example, the Navajo language has two terms, ***dził*** and ***yílk'id***, which have meanings that fairly closely parallel the meanings of *mountain* and *hill*, respectively in English. However, the Navajo term ***dził*** is applied to large features covered with pine forest, even if these are called plateau or mesa in English. For features composed of exposed bedrock, the Navajo language uses neither ***dził*** and ***yílk'id***, but employs compound terms starting with ***tsé***, the Navajo term for rock, followed by terms for shape or posture, such as standing, sitting, or lying.

We have similar data for other landscape domains, and again find that the details of the categories are different for the different languages. We also have data for other languages that confirm that categories for landscape features and components do not 'line up' across languages.

2.3 Landform Extraction

As noted above, digital elevation models (DEMs) are discrete approximations to elevation fields. Landforms such as hills or valleys, on the other hand, are regions or parts of the Earth's surface that have a coherence of shape, earth surface processes, or other factors not directly represented in the DEM (Mark and Sinha, 2006). There is interplay between feature extraction and delimitation, quantitative description of the features, and classification of the features into meaningful categories. There is circularity here: one can only compute the size and shape of topography within some region, yet the detection and delimitation of a form depends on the computed size and shape. It is likely that this issue can only be overcome by iterative solutions. Detect a candidate feature, then delimit it, then determine its shape, then classify it, and then use the characteristics of the category to make a better determination of boundaries, re-parameterize, re-classify, until there is a stable solution. The phase where we "use the characteristics of the category" likely will be different for different languages. Sinha's (2007) dissertation provides a good start on the problem for eminences and for terms in English.

3. Summary

Intelligent multilingual landform extraction and classification presents a number of interesting research challenges. Bottom-up DEM-based measures of local form may be useful in some situations, but more global or top-down approaches are likely needed if the goal is to detect, delimit, and classify landforms that correspond to the categories used by earth scientists or by the general information-retrieving public.

Acknowledgements

This work is being done in collaboration with Gaurav Sinha (Ohio University), Andrew Turk (Murdoch University, Perth, Australia), David Stea (Texas State, San Marcos), the Yindjibarndi people (Roebourne, Australia), and the Navajo Nation and people. Discussions with Boyan Broderic and Werner Kuhn have been valuable. The research is supported in part by NSF Grants BCS-0423075 and BCS-0423023 and by an award from Microsoft Research.

References

Evans, I. S., 1972. General geomorphometry, derivatives of altitude, and descriptive statistics. In: R.J. Chorley, Editor, Spatial Analysis in Geomorphology, Methuen, London, pp. 17–90.

Mark, D. M., 1975. Geomorphometric parameters: a review and evaluation. Geografiska Annaler, 57A, 165-77.

Mark, D. M., 1993. Toward a Theoretical Framework for Geographic Entity Types. In Frank, A. U., and Campari, I, editors, Spatial Information Theory: A Theoretical Basis for GIS, Berlin: Springer-Verlag, Lecture Notes in Computer Sciences No. 716, pp. 270-283.

Mark, D. M., and Sinha, G., 2006. Ontology of Landforms: Delimitation and Classification of Topographic Eminences. In M. Raubal, H. Miller, A. Frank, and M. Goodchild, Eds. Geographic Information Science - Fourth International Conference, GIScience 2006, Münster, Germany, September 2006, Extended Abstracts. IfGI prints 28. Verlag Natur & Wissenschaft, Solingen, Germany, pp. 129-132.

Mark D. M., and Smith, B., 2004. A Science of Topography: From Qualitative Ontology to Digital representations. Chapter 3 in Michael P. Bishop and John F. Shroder, editors Geographic Information Science and Mountain Geomorphology Chichester, England: Springer-Praxis, pp. 75-100.

Mark, D. M., and Turk, A. G., 2003. Landscape Categories in Yindjibarndi: Ontology, Environment, and Language. In Kuhn, W., Worboys, M., and Timpf, S., Editors, Spatial Information Theory: Foundations of Geographic Information Science, Berlin: Springer-Verlag, Lecture Notes in Computer Science No. 2825, pp. 31-49.

Mark, D. M., Turk, A. G., and Stea, D., 2007. Progress on Yindjibarndi Ethnophysiography. In Winter, S., Duckham, M., Kulik, L., Kuipers, A., (editors) Spatial Information Theory. Lecture Notes in Computer Science No. 4736, pp. 1-19.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., and Oltramari, A., 2003. WonderWeb DeliverableD18, Ontology Library (final). Laboratory For Applied Ontology, http://www.loa-cnr.it/Papers/D18.pdf.

Peuquet, D.M., Smith, B. and Brogaard, B., 1998. The Ontology of Fields. Varenius Project Technical Report, (Santa Barbara, CA: National Center for Geographic Information and Analysis (NCGIA).

Sinha, G., 2007. Delineation, Characterization and Classification of Topographic Eminences. Unpublished Ph.D. dissertation, Department of Geography, University at Buffalo.

Sinha, G., and Mark, D. M., 2010. Extraction and Database Modeling of Topographic Eminences. Cartographica, under consideration.

Smith, B., and Mark, D. M., 2003. Do mountains exist? Towards an ontology of landforms. Environment and Planning B: Planning and Design, 30(3), 411-427.

Turk, A. G., and Mark, D. M., 2008. Illustrated Dictionary of Yindjibarndi Landscape Terms. Informal publication, Murdoch Univerity (Australia), October 2008.

Voegelin, C. F., and Voegelin, F. M., 1957. Hopi Domains: A Lexical Approach to the Problem of Selection. Memoir 14 of the International Journal of American Linguistics. Baltimore: Waverly Press, Inc.