Modal Inference

Whitman Richards

Mass. Inst. of Technology
CSAIL 32-364, Cambridge MA. 02139
wrichards@mit.edu

Abstract

Evidence is presented that properties of objects, or events, or man-made artifacts in the natural world are clustered in a multidimensional space. Hence observations made by different sense modalities can be highly correlated. Such clustering favors perceptual inference, increasing its robustness, as well as supporting generalizations.

The Claim

When we, or intelligent agents in general, interact with objects or other creatures, our choices for actions are typically based on the perception and recognition of the current state of the world. This is an inferential process, involving the analysis of very limited sense data. Why, then, are our percepts so reliable and successful?

A Darwinian explanation would be that creatures who are successful in their environments have evolved powerful inferential machinery through natural selection. However, another explanation, compatible with Darwinian evolution, is that objects, events, behaviors -- even cognitive models and social conventions – appear in clusters in a multidimensional "property" space. In other words, given one inferred property (i.e. based on observations of one sensor modality), there is a high probability of inferring several other properties though other sense modalities. Such clusters of correlated properties are called "Natural Modes" (Richards & Bobick, 1988.) If a world hosts such modal clusters, then these should be the focus for understanding successful survival.

Definition: A Natural Mode is a cluster of lawful regularities, where the observation of one regularity is highly predictive of a set of other regularities. (See Jepson & Richards (1993) for a more formal statement.)

As will be seen shortly, the lawful aspect of modal properties is important, because one can then generalize from one to other objects within a modal cluster.

Simple Example

If the modal inference hypothesis holds, then Nature's generative processes must lead to correlated properties, which then can be inferred using quite different modes of observation. Consider viewing and grasping a metal object. The visual observations include Lambertian and distinctive specular components; the shape appears solid (rather than undergoing fluid distortion); when grasped, the surface is very hard and often smooth; it may also feel cool; if hit with a hard stick or knuckle, there is a characteristic metallic sound due to internal friction (Richards & Wildes, 1988.) A cotton ball would have a quite different set of observations; a puddle of water another. Recently, Michael Coen (AAAI 05) has used the modal correlates of speech (sound and lip movements) to categorize the vowels using a self-supervised Hebbian learning algorithm. Such bootstrapping of information across modalities can be achieved only if the underlying generative processes are correlated. Regardless of whether or not we understand the essence of these generative processes, powerful and reliable inference can be achieved by observations of different modalities that are independent. More, however, can be inferred if the properties observed obey allometric rules.

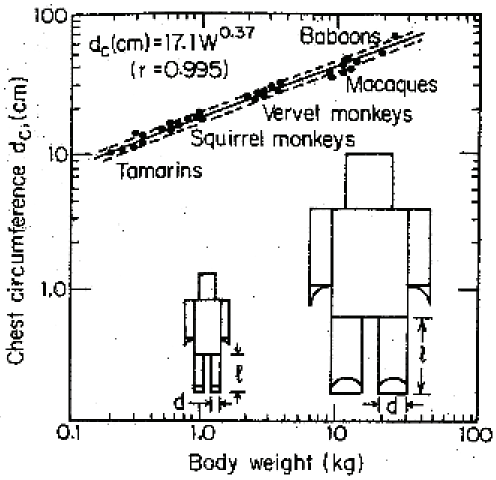

Fig. 1. Example of an allometric relation (McMahon, 1973)

Allometry

Allometry is the generalization of simple power laws that describe different rates of growth or form for parts of a body relative to the body as a whole (Thompson, 1917;

Huxley 1932.) In Fig. 1 is one example. Shown are the relation between for chest size and body weight (upper curve) as well as the diameter to length ratio of limbs needed to support body weight. The intuition is that thicker legs are needed to support large masses. A similar relation applies to other structures, such as trees. Seedlings have very slender trunks relative to the size of the arborized body, whereas Sequoias have relatively thick trunks.

Fig. 2 presents another example. Here leg length is related to stride frequency. To first order, the relation is approximated by an inverted pendulum, such as a wooden toddler might use.

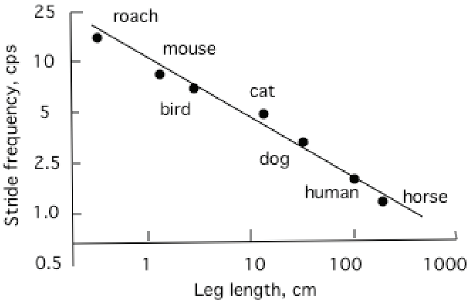

Fig. 2. leg length vs stride frequency (Richards & Bobick, 1988)

To create a mode, we link several such allometric relations, as shown in Fig. 3 below. Here, each graph depicts a (power law) relation of a body part to leg length for a large class of terrestrial animals. Body and "hand" size increase with leg length (upper left); animals with long legs "lumber"; those with shorter legs "scamper (lower left.) If body size increases, so will the vocal tract length with a decrease in the pitch of emitted sounds. As shown by the dashed lines, given one type of observation in the context, several other properties may be inferred. In this example, the observation modes include visual shape and aspect, rate of movement, and speech and footfall sounds.

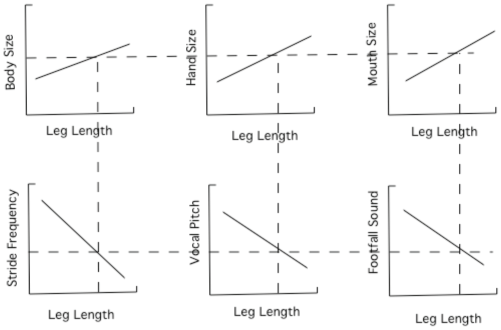

Fig. 3.Schematic of size lawful relations to leg length (adapted from McMahon, 1973, 1975.)

Modal Evolution

T. H. Huxley (Maynard-Smith et al, 1985) once commented that "there appears to be" predetermined lines of modification that lead natural selection to "produce varieties of a limited number and kind" for each species. Modal regularities in designs would support this conjecture. Especially significant is when creatures, such as Man, have the knowledge and capability of modifying and reforming their environments, passing this knowledge on to successive generations. Language and culture then become significant forces in evolutionary development (Waddington, 1959; Cavalli-Sforza, L and Feldman. M., 1981) The invention of hand tools (which also can be shown to have occurred in steps of punctuated modal evolution – see Fig 4), enabled humans to create many useful artifacts that have modal properties (e.g. dwellings, writing implements, means of transportation, social conventions….)

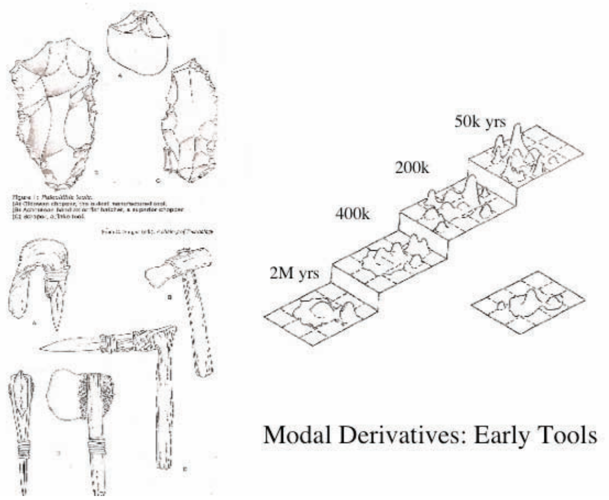

Fig. 4.Early tool development showing the variation and increase in differentiation over time. The two dimensions for each step represent two different measurements. The four steps show different times, as indicated, The lower right shows tool measurements for a modern Australian aboriginal tribe. (From G. L. Isaac, 1976.)

Fig. 5 offers a sketch of the development of some important artifacts that have led to a branching hierarchy of sub-modes. (Their root is in the modal structure of the human being.) The graphical form resembles proposals for the evolution of language. Which might also be considered a modal endeavor.

Elsewhere, we have shown how the basic structure of modal relations can support reliable perceptual inference and categorization (Bobick , 1986; Richards, Jepson and Feldman, 1992; Feldman, 1997).

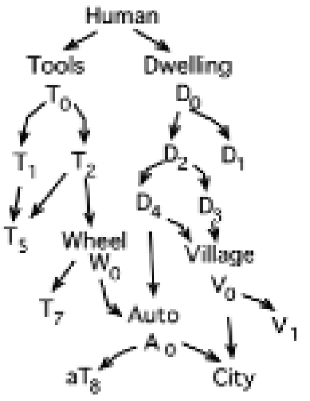

Fig 5. Hierarchical sub-mode evolution derived from Human modal structure

An intriguing question is to what degree the space of our models for physical and social phenomena also have a modal character – at least at our own space-time scale. The answer may clarify Wigner's commentary about why our theories of Natural phenomena are typically so successful.

References

Bobick, A., Richards, W. 1986, Classifying objects from visual information. MIT- AI Lab Memo 879. pp1-42.

Cavalli-Sforza, L and Feldman. M. 1981, *Cultural transmission and evolution: a quantitiative approach*. Princeton Univ. Press.

Coen, M. H. 2005. Cross-modal clustering. Proc. 20th Conference on Artificial Intelligence (AAAI'05).

Feldman, J. 1997. The structure of perceptual categories. Jrl. Math. Psychology **41**, 145 – 170.

Huxley, J. 1932. *Problems of Relative Growth* (2nd Ed, 1972) Dover.

Isaac, G. L. 1981 Emergence of human behavior patterns. Phil. Trans. R. Soc. Lond. B **292**, 177 – 188. (see 1976 ref.)

Jepson, A. and Richards, W. 1993, What makes a good feature? In: *Spatial Vision in Humans and Robots*, L. Harris & M. Jenkin (eds.) Cambridge Univ. Press.

Jepson, A., Richards, W. and Knill, D. 1996. Modal structure and reliable inference. In: *Perception a Bayesian Inference*. D. Knill and W. Richards (Eds), Cambridge Univ. Press.

Maynard-Smith, J., Burian, R., Kaufmann, S., Alberch, P, Campbell, J., Goodwin, B., Lande, R., Raup, D., Wolpert, L. 1985, Developmental constraints on evolution. Quart. Rev. of Biology **60**, 265-287.

Marr, D. 1970. A theory for cerebral neocortex. Proc. Roy. Soc. Lond. B, **176**, 161-234.

McMahon, T. A. 1973. Size and shape in biology. Science 179, 1201 – 1204.

McMahon, T. A. 1975. Using body size to understand the structural design of animals. J. Applied Physiology **39**, 619-627.

Niyogi, P. 2006, *The computational nature of language learning and evolution.* MIT Press.

Richards, W., and Bobick, A. 1988. Playing Twenty Questions with Nature. In: Computational Processes in Human Vision: An interdisciplinary perspective, ed. Z. Pylyshyn. Ablex, Norwood, NJ.

Richards, W., Jepson, A. Feldman, J. 1992, From features to perceptual categories. Proc. British Machine Vision Conference, Leeds, September 1992, pp. 99-108.

Richards, W. and Wildes, R. 1988. Recovery of Material Properties. In: Natural Computation, W. Richards (Ed.), MIT Press.

Thompson, D'arcy 1917. *On Growth and Form* (reprinted 1968) Cambridge Univ. Press.

Waddington, C. H. 1959, Evolutionary systems – animal and human. Nature **183**, 1634 – 1638.

Wigner, E. 1960, The unreasonable effectiveness of mathematics in the Natural Sciences. Comm. Pure & Applied Math, **13**, 1 – 14.

Acknowledgement: Supported in part by AFOSR MURI initiative FA9550-05-1-0321.