Poetics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Poetics

journal homepage: www.elsevier.com/locate/poetic

Research note

A formal approach to meaning

John Levi Martin[a,*], Monica Lee[b]

[a] Department of Sociology, University of Chicago, 1126 E 59th Street, Chicago, IL 60637, United States
[b] Facebook, Inc., United States

ARTICLE INFO

Keywords:
Meaning
Information theory
Concept analysis
Entropy

ABSTRACT

Building on the overlap between notions of duality and the principles of Formal Concept Analysis, we propose a way of quantifying the degree of meaningfulness of one aspect of a data set for defining another. Drawing on ideas from information theory, some familiar and some less familiar, we demonstrate not only how to maximize the meaningfulness in a dataset, but how to translate across datasets.

1. Introduction

1.1. Explanation: meaning versus outcomes

In much of sociology, we attempt to explain *outcomes*, which, aside from often being a fool's errand (since not everything that happened *had* to happen), tends to dovetail with issues of inference that are equally problematic. For example, we do not seem to be in a strong position to use conventional sociological methods to demonstrate that a certain event (a revolution, say) was caused by a certain factor or set of factors unless we can make the claim that there is a set of cases, real or imaginary, to which our argument could be extended such that every member of that set would also undergo the outcome, were all factors, measured or not, identical.

Quite different is when we link our observations in terms of *meanings*. This is often associated with a lack of interest in issues of inference: to say that the meaning of the hump on Moby Dick's back is a pyramid is not to say that all whales' humps mean pyramids, nor even that in nineteenth century American literature, whales' humps are more likely to mean pyramids than are other humps. While meaningful explanations can work without supporting inferential statements—they can be about singularities—outcome explanations cannot work without supporting meaningful ones. And this is because our idea of explanation is one that entails meaningfulness. If we were to claim to explain why one respondent in a survey supports gun regulation, but we were to do so by pointing out that her vector of response on 12 other questions was [9,4,1,29.34,na,0,0,1,9383,na,0], and that this uniquely predicted favoring regulation, then, even if this pattern had strong out-of-sample validity, if we could say nothing further as to why *these* values, on *these* times, *reasonably* could be expected to predict support, no one will accept that we have done any explaining. In some cases, however, we may combine the approach to the explanation of meanings with the explanation of data regularities, because we treat the vector responses not as predictors but as links in a web of meaningful relations.

Let us give an example. In the men's sale aisle at Shoes For Less are pink leather high tops. "No wonder," we think. We would, perhaps, expect high tops to be canvas, not leather; we would expect leather to be black, not pink. These clumps of regularity in the space of possibility (in this case, a feature space) arise not because of causal constraints in the production process, but because of what these attributes mean. For this reason, we can use such distributions to answer questions of meaning.

"You are a pig!" someone says. What does this mean? We know from repeated usage that this metaphor is used to transfer certain

* Corresponding author.
E-mail address: jlmartin@uchicago.edu (J.L. Martin).

https://doi.org/10.1016/j.poetic.2018.01.002
Received 25 June 2017; Received in revised form 18 January 2018; Accepted 18 January 2018

attributes of the pig—most importantly, its inattention to the presence of filth on its body—to the tenor (in this case, you). But what does, say, a pig mean in a work of literature? To take one very simple example, what is the meaning of the pigginess of a character in the (extremely popular) children's book by Richard Scarry, *What Do People Do All Day?*

1.2. Our approach

We here are inspired by Breiger's classic (1974) work, and the way in which his conception of duality (also 2000) dovetails with conceptions of meaning developed first by Marx and Engels (1976 [1845–6]). To Marx and Engels, the meaning of some abstract idea could only be understood when we know how it arose from a pattern of real relations between sets of persons. In a simpler version, Bourdieu ([1979] 1984) proposed that the meaning of some cultural object or practice could only be understood when we know the profile of the types of person who appreciate it or who participate in it. In Bourdieu's work, a large data matrix with persons as rows and cultural practices as columns allows us to understand the meaning of the columns by using the distribution across rows.

Breiger's (1974, 2000) conception of this fundamental issue is one that highlights the fact that our attempts to understand meaning can be envisioned as walks from one set to another, and back again. It is akin to the formalization by White (1963b) of the analysis of Lévi-Strauss ([1949] 1969), and extended by Boorman and White (1976) and Breiger and Pattison (1986), as well as by Lee and Martin (this issue)—we use matrices to reproduce walks through a structure. However, Lee and Martin (2015) concentrate on how to ask new types of questions about the relation between social and cultural meanings; here we wish to consider whether existing quantifications can guide us as to *which* sets of relations to traverse. With this conception, we make use of existing statistics, most of which are well known, to formalize an approach to the study of meaning.

1.3. A formal approach

We here build on well-known approaches in formal concept analysis and information theory to make a simple suggestion for how to formalize notions of meaning. The essence of our proposal is to generalize the approach to meaning implicit in formal concept analysis using one particular approach to information theory.

We consider "meaning" an undefinable term, like "set." Accordingly, our goal here is not to develop formal procedures based on a particular definition of meaning, or to attempt to replace less formal definitions with a formal one. Instead, we begin with what we believe to be the core notions of "meaning," and develop this in a formal manner, hoping to develop meaningful ways of quantifying meaning.

The term "meaning" is ambiguous, as it may refer to something's *significance*, or to its *sense*. It is the latter with which we begin. In this sense of meaning (the meaning of meaning *as* sense), the meaning of one thing is something else. Not all references are meaning-references, and we cannot precisely define meaning-references, but the centrality of such reference allows us to conceptualize meaning as a connection of one thing to another. The meaning of x is then A, and hence meaning here is the x-A relation. More generally, there is no reason why the meaning of something cannot be a one-to-many relation. Thus the meaning of x may not simply be A, but A and B. It would seem that adding B on, and therefore increasing specificity in our predication of x, also increases the "meaningfulness" of the predication. An approach to such meaning relations has already been developed using algebraic logic, and we begin here.

2. The meaning of meaning

2.1. Intension and extension

In logic, a generalizing concept, such as a class concept (e.g., *tree*) may be said to have an *extension*—the set of individuals that it subsumes. But it also may have an *intension*—a web of connections to other concepts. A key question in the logical analysis of meaning has historically been whether these two meanings of meaning were fundamentally the same, such that intension could be defined as extension (a move appealing to many nominalists). Most logicians concluded that the two could not be the same, as they varied inversely—the greater the extension of a concept, the less its intension, as it says the least about any item (Kant, 1988 [1800]: 101f).[1] To say that an item is a "tree" is less meaningful than to say it is "a deciduous tree" which is to say less than to say that it is an "oak tree."

It is this vision of the meaning of concepts that brings us to formal concept analysis (Ganter & Wille, 1999). The exemplary data for such concept analysis is a set of individuals (or species) and a set of characteristics (or "marks"). We here will consider most generally a relation R between two sets, A and B, with generic members a and b respectively, which may or may not be disjoint. Because the first set is taken to be the set of objects, and the second set the set of characteristics, we define two mappings, which will be called "intent" and "extent" respectively, the first $A \rightarrow B$, and the second $B \rightarrow A$, as follows:

$$\Gamma(a) = \{b \in B | aRb\} \tag{1}$$

[1] Note that the term "intension," supplied by William Hamilton—and rued by Peirce (1965: 2.393) who preferred the term "comprehension," used by the Port Royalists—is anachronistically used in the translation of Kant for *Inhalt*; likewise *extension* for *Umfang*.