LexiClone
lexical cloning system

The Role and Meaning of Predicative and Non-predicative Definitions in the Search for Information.

Ilya S Geller
LexiClone Inc.

Borrowings and Adaptations.

First of all, the author would like to clarify the meanings of the terms "predicative" and "non-predicative". As is well known, *praedicatum* in Late Latin means "what has been said (previously)". In Aristotelian and subsequent forms of traditional logic [8,9,11,12,13,14,15,16] a predicate was understood to be one (the one in which something is said about the subject of speech) of the two terms for the judgment of a subject. In his treatment the author counts as predicative any definition of a subject or object in which something is said about an observable subject or object as it changes. In addition, the author proposes as the sole measure of change the movement of a subject or object with acceleration [5,6,7,17]: if a subject or object is immobile or moves evenly, it cannot be observed and, consequently, cannot be defined predicatively [1].
More: as is well known, Bertrand Russell introduced the notion of a "non-predicative" definition, in which what is to be defined is brought in through its relation to a class of which it is an element. For example: "the set of all sets that are not elements of themselves". It is said that the use of "non-predicative" definitions leads to paradoxes, so they should be dealt with carefully.
The author adopts Russell's definition [27,28,29,30,31], but in a new mode: one counts as non-predicative any definition of a subject or object in which something is said about a subject or object that is unchanging. For example, according to the author there are no closed sets in the world of change - a set can strive to be the set of all sets, but no more.

The Text.

A text is made up of words. But what is a word? First of all, a word is made up of letters, which are, in practice, meaningless if separate. And yet a word, as the joining together of several letters, already, beyond any doubt, has a certain meaning. But the existence of synonymy makes the meaning of words, taken separately, vague and lacking in concreteness and separate words are declared to be non-predicative definitions. For example, the word “red”, taken by itself, can mean anything: beginning with a colour and ending with a pejorative name for a Communist. In order to understand the “true” meaning of a word one must first identify in what minimal lexical construction of speech and in what slang a given word is being used; where:
I. A minimal lexical construction of speech is a predicative definition: the articulation of three words, relating to three parts of speech - substantives, verbs and adjectives - in the context of a sentence. All other parts of speech, with the exception of prepositions and interjections, can be, in some way or other, taken to be substantives, verbs, and adjectives, where:
1. A substantive has the meaning of the abstract Name of certain points of accumulation[1];
2. A verb defines the abstract Name of an action;
3. An adjective is the abstract Name describing points of accumulation in the process of change.
II. A slang is an aggregate of predicative definitions used with a strictly specialized meaning, particular to each and every type of human activity. [2,3,4,22,26,33,36]
This triad of non-predicative definitions is indispensable to defining the subjective evaluation of a fact, when faced with the possibility and the need to include the objects and subjects of the fact

[1] A point of accumulation is the boundary point of the set M - the point x of the topological space X>M, of which any vicinity contains an indefinite number of points of the set M. Everything – all things, animals, humans, etc. – are “accumulation points”. [8,9,11,12,13,14,15,16,35,36,38].

PDF created with FinePrint pdfFactory trial version www.pdffactory.com

within the certain context. In other words, a person has to evaluate a sandwich from all sides: he has to understand that it's a sandwich, and to decide whose it is, whether he should eat it or not, whether it's fresh and tasty, etc.
Moreover, the presence of at least one predicative definition is absolutely necessary and sufficient for the creation of a sentence, even if it's missing one or more words from the substantive/verb/adjective triad. Such a word or words can be reconstructed on the basis of the context and subtext of the predicative definition; where:
1. The context consists of those predicative definitions where a substantive is used as the abstract Name of points of accumulation and abstractions;
2. The subtext consists of those predicative definitions where pronouns and interjections are used as the abstract Name of points of accumulation and abstractions.
For example, having said the word “unfresh”, one can reconstruct the words “sandwich” and “exists” if we know in what context and subtext the word “unfresh” appears. And if we don't know the context and subtext of a given predicative definition, then the word “unfresh” can be used with, for example, the words “fish” and “smells”. Only a text, being a collection of predicative definitions grouped together in meaningful sentences, can provide, more or less identically, the context and subtext of every one of these predicative definitions. That is, a text is considered to be completed in so far as its context and subject are, more or less identically, defined.

Processing the Text.

The task comes down to extracting all the predicative definitions from every sentence of the text, and then counting how many times each one occurs in the text. Such a collection is termed a summary; the number of times each predicative definition occurs in the text is referred to as its weight. (NLP uses the notion of predicative definitions not counting weights of them.) A summary, being an ordered list of triads, is susceptible to rapid processing by computer. The margin of error in the cloning is lessened in proportion to the amount and size of the texts being used.

Examples of Summary.

The entire summary of George Bernard Shaw, created on the basis of his books as found on the Internet at the URL http://promo.net/pg/ , contains a little over 320,000 triads occurring more than once. The first triad - it-be-in - occurs 4 755 times; the second in order of frequency - i-be-in - occurs 2 534 times [18,19,20].
Similarly, at the URL http://lexiclone.com/SummarySample_Fyodor_Dostoevsky.htm , the reader can see an extract from Fyodor Dostoevsky's summary (a part of which is reproduced below), created on the basis of his book The Brothers Karamazov (the numbers to the right represent the frequency of each triad-phrase's occurrence in the text):

it - be - in : 1 466
i - be - in : 1 347
it - have - in : 996
you - be - in : 936
you - be - your : 798
i - have - in : 664
all - be - in : 657
it - will - in : 535
my - be - in : 496
all - have - in : 473

Clearly, it is subtext that dominates in Dostoevsky's text. In another summary - that of our Patent #6.199.067 - it is context that dominates [32]:

one - say - least : 1 447
segment - say - least : 1 124
datum - item - plural : 1 025
system - say - remote : 950
datum - say - plural : 888
computer - say - remote : 845
datum - item - linguistic : 845
system - say - least : 844
computer - say - least : 818

PDF created with FinePrint pdfFactory trial version www.pdffactory.com

one - say - remote : 805

It appears that the preponderance of subtext can be explained by the fact that certain texts concentrate on the Ethical component in the process of a person's becoming whole[2] - on the question, what will happen if certain points of accumulation are included in the vicinity of a given? - rather than on the Aesthetic component. On the other hand, texts of the kind we might call "technical" are primarily concerned with the Aesthetic component, examining not the consequences of an action but rather its mechanism, and therefore context predominates in them. The summary of this article is:

weight - be - in : 592
weight - be - summary : 436
text - be - in : 418
it - be - in : 306
text - be - summary : 265
weight - weight - in : 214
i - be - in : 210
weight - weight - summary : 194
weight - say - in : 182
say - weight - in : 182
triad - be - in : 176
each - be - in : 176
all - be - in : 169
say - say - remote : 168
say - say - in : 164
text - weight - in : 158
phrase - be - in : 156
text - weight - summary : 134
one - be - in : 134
say - say - plural : 124
text - say - in : 116
say - be - least : 114
each - be - summary : 114
triad - be - summary : 109
triad - weight - summary : 61

To show summaries in their entirety would be impossible because of their extremely large (up to 3M) size [34].

The Search Engine.

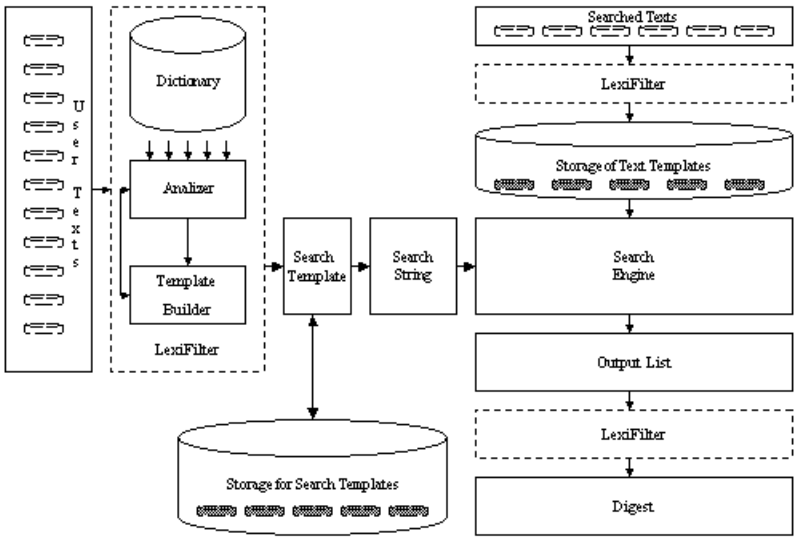
The author, as an immigrant in the lair of unbridled capitalism with its limitless opportunities to "make" money, naturally put the theory into practice at once and created a search engine for finding textual information in electronic form: the program is demonstrated on the Internet at the URL http://www.lexiclone.com/Products.htm

The system works as shown in the following diagram:

[2] The author assumed that a point of accumulation (an open set) always strives to become a/the material point (a/the closed set); and that this striving is the motivation for the universe to "spin" around, to change: the point x of the topological space X>M strives to include (for the sake of closeness) some other points of accumulation in its neighborhood. But after becoming complete-closed a set (a point of accumulation) is to lose all its qualities that make it unique and distinct. Here, the author sees the well-known Russell's paradox: how to distinguish something that has exactly the same quality? In other words: if something is "red", how to distinguish this "red" from all other "reds" (a complete-closed set does not change and cannot be predicatively defined)? [35,36,37,38]

PDF created with FinePrint pdfFactory trial version www.pdffactory.com

Figure 1



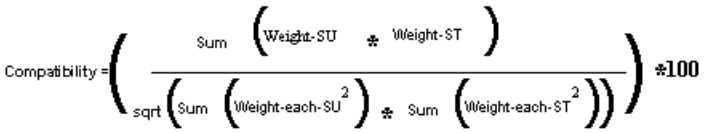
The program, called UniSearch, does the following:
It fixates the searcher's summary, as a network of predicative definitions within limits; compares it with the summaries of authors, taken within limits, for a certain collection of texts. The result of the operation of the program consists of several (sufficiently few) sentences or predicative definitions, distilled out of the collection of texts, containing the desired information and called "digests".
These are some examples of searches' results: the author asked clones *How should we prosecute terrorists, murderers who kill innocent people? What should we do with them? Should we kill them in return?* The clone of Plato said: *How should we answer him?* The clone of the Bible said: (4.0% Ephesians) *For we are his workmanship, created in Christ Jesus unto good works, which God hath before ordained that we should walk in them*. The clone of Mohamed said: *Thee do we serve and Thee do we beseech for help*. The clone of Bernard Shaw (20th century) answered: *Who are we that we should judge them?* Fedor Dostoevsky said - *Well, what can we do?*
For the normal functioning of the program it is desirable, although not necessary, to create a summary of the user. Then,
1. The user puts in a search text, of any size;
2. On the basis of that text, the program selects those texts in the collection that were created by authors whose own summaries are closest to the summary of the person seeking the information, as extracted from the search text.

Compatibility.
The choice of one or another text is made through the medium of a standard formula, which is

PDF created with FinePrint pdfFactory trial version www.pdffactory.com

called a “controller” in Reinforcement Learning, and later named Compatibility by the author:

Figure 2



where:
Weight-SU is the weight of the triads that are common to the summaries of the searcher and of the text,
Weight-ST is the weight of the predicative definitions that are common to the user’s summary and the summary of the text,
Weight-each-SU is the weight of each triad in the user’s summary,
Weight-each-ST is the weight of each predicative definition in the summary of the text [23].

References.

[1] Aristotle, A new Aristotle reader, Princeton University Press, Princeton, 1995
[2] A. Ayer, The Foundations of Empirical Knowledge, Macmillan, 1940
[3] A. Ayer, Logical Positivism, The Free Press, Glencoe, Illinois, 1960
[4] A. Ayer, Language, Truth and Logic, Dover Publications, Inc., NY, 1952
[5] N. Bohr, Atomic Physics and Human Perception, Moscow, 1962
[6] N. Bohr, Can Quantum Mechanical Description of Physical Reality Be Considered Complete? Moscow, 1992
[7] Bohr N., Continuity, Determinism and Reality”, Moscow, 1992
[8] F.H.Bradley, The Principles of Logic, Oxford University Press, London, 1883
[9] F.H.Bradley, Appearance and Reality, Oxford at the Clarendon Press, 1959
[10] Brill E., Susan Dumais and Michele Banko. An Analysis of the AskMSR Question-Answering System
[11] G. Hegel, Science of Logic, Vols.1-3, Moscow, 1972
[12] G. Hegel, Phenomenology of Mind, George Allen&Unwin Ltd., London
[13] G. Hegel, Lectures on the History of Philosophy, Moscow, 1932
[14] G. Hegel, On Art, Religion, Philosophy, Harper Tourchbooks, Harper&Row, Publishers, NY, 1980
[15] G.Hegel, The Encyclopedia Logic, Hacket Publishing Company, Inc., Indianapolis, 1991
[16] G. Hegel, The Essential Writings, Harper Torchbooks, Harper&Row, Publisher, NY
[17] W. Heisenberg, Problems of Philosophy, Moscow, 1953
[18] D. Hume, Dialogues Concerning Natural Religion, Hacket Publishing Company, Cambridge, 1984
[19] D. Hume, Enquiries concerning the human understanding and concerning the principles of morals; Oxford, Clarendon P., 1969
[20] D. Hume, Treatise of Human Nature, Oxford at the Clarendon Press, 1951
[21] R. Geyer, Alienation theories; Oxford, 1980
[22] W. James, Pragmatism, New York, 1902
[23] H. Lebesque, About Measurement of Measure, Moscow, 1969
[24] L. Peshkin, Reinforcement Learning by Policy Search (dissertation)
[25] L. Peshkin, Edwin D. de Jong. Context-based policy search: transfer of experience across problems
[26] Pierce, Collective papers. Vol.5, Cambridge, 1960
[27] B. Russell, Scientific Method in Philosophy, Oxford, 1914; Moscow, 1988
[28] B. Russell, Introduction to Mathematical Philosophy, London, 1953; Moscow,

PDF created with FinePrint pdfFactory trial version www.pdffactory.com

1989
[29] B. Russell, Principles of mathematics; Bradford & Dickens, London, 1950
[30] B. Russell, The Problems of Philosophy, Hacket Publishing Company, Inc., Indianapolis, 1996
[31] B. Russell, My Philosophical Development, Ruskin House, George&Unwin Ltd, London, 1959
[32] G. Ryle, Dilemmas, Cambridge at the University Press, 1954
[33] D. Sack, Human Territoriality; Cambridge University Press, Cambridge, 1986
[34] P. Turney, Mining the Web for Lexical Knowledge to Improve Keyphrase Extraction: Learning from Labeled and Unlabeled Data

Some Extra References.

[35] P. Aleksandrov, "An introduction in the theory of sets and general topology", Moscow, 1977
[36] I. Geller, New Mechanics: The Foundation, http://lexiclone.com/NEWONE1_BODY.html
[37] I. Geller, In Search of a Living Language, http://lexiclone.com/insearchofalivinglanguage.htm
[38] G. Cantor, Works on the theory of sets, Moscow, 1985

PDF created with FinePrint pdfFactory trial version www.pdffactory.com