Computability by “Distinctive” Syllogistic for Ontologies and Semantics

Daniele Ingrassia[1]**, Ferdinando Cavaliere**[1]

Abstract. This paper has two main purposes.

The first purpose is to show how a particular logical structure, called "Distinctive" Syllogistic, closer to the natural language than the classic one, but equally rigorous, is able to describe in a simple way the relations between concepts and the deductions that they can build.

The logical structure, based on five possible relationships between two concepts, and the quantifier Y ("partial") can be read in both key: bivalent or fuzzy. This predicate calculus is applied in different fields of human knowledge, and reveals recurring patterns in the organization of the concepts, almost an "a priori" of our categories and mental attitudes. The precision of its expressions, the elimination of certain ambiguities of language, is accompanied by an elasticity that make it suitable to bring together natural and artificial cognitive systems. Its applications in scientific fields such as those in the humanities are exemplified in a table of comparison.

The second purpose is to adopt the Distinctive Syllogistic as a tool of organization, research and translation of concepts in all areas of communication and artificial intelligence, in particular in the field of semantic search engines, construction of databases, dictionaries, translators.

1 INTRODUCTION: PREDICATES OF “DISTINCTIVE” SYLLOGISMS AND NATURAL LANGUAGE

Taking an ordered pair of term/sets *ba*, we can use categorical predications to express the various possible cases. That is, we have:

- universal affirmative “**every b is a**”, in syllogistic symbols: **Aba** (e.g. the dog is a mammal);
- universal negative “**no b is a**“: **Eba** (e.g. the dog is not a cat); [1]
- *distinctive* or *partial* particular “**only some b is a**”: **Yba** (e.g. only some mammal is a dog);

The quantifier Y represents the intuitive natural language *some* and not the existential *some* of classical predicate logic. The latter, symbolized by **Iba**, means ‘at least some, perhaps all’, not excluding the universal quantifier, whereas the former stands for ‘only some’ or ‘at least one but not all’; see [3].

The Simple Distinctive Calculus (named D3 or D6), which can be developed on this categorical predications, constitute a Syllogistic deductive system and is the object of another study by one of the present authors [4].

From a set-theoretical and diagrammatical point of view (Venn , Gergonne, see [9]), there are 5 distinct cases of possible relations between two sets: I. identity or equivalence, II. proper inclusion of the first in the second, III. proper inclusion of the second in the first, IV. mutual proper (non-null and non-exhaustive) intersection, V. exclusion or incompatibility. As shown in figure 1, owing to the three categoricals, we obtain an exhaustive tri-partition of the five cases, something which would not be possible using the traditional partials: for example the cases I or II can validate Aba as well as Iba).

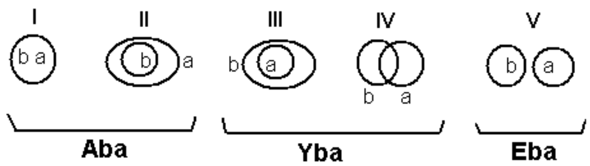

Figure 1. Exhaustive tri-partition of the five cases

From a semantic point of view, it enables us to express intermediate or “fuzzy” concepts, essential for a proper account of linguistic communication. The five cases lend logical support to the common notions of similarity (affinity) and difference. These may form two inversely proportional scales in the sense that the more two terms are similar, the less they are different, and vice versa. See figure 2.

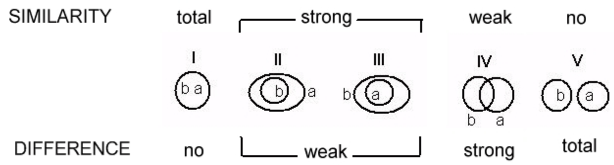

Figure 2. Similarity and difference

2 SEMANTIC APPLICATIONS OF THE 5 “DISTINCTIVE” RELATIONS

The three categoricals are insufficient to pick out uniquely the five situation classes to which they can refer. In order to describe uniquely each of the five cases in predicative form, we can add, by means of the conjunction *and*, to the three predications over the pair *ba* those generated by the inverted predicative pair *ab* (see column c in figure 3).

[1] Circolo Matematico Cesenate, Cesena (FC), Italy. Email: {ingrassiada, cavaliere.ferdinando}@gmail.com

a) SET RAPRESENTATIONS	b) Gergonne's NOTATION	c) CONJUNCTION OF INVERTED CATEGORICALS	d) QUANTIFICATION OF PREDICATE	e) LIBRARIANSHIP subjets' cataloging	f) SEMANTIC - DICTIONARY	g) RETHORIC THROPHES
I	b I a	Aba * Aab	AbAa	Use For all-all relation	synonymous	Definitio - Periphrasis
II	b ⊂ a	Aba * Yab	AbYa	Broader Term all-some rel.	hyponymous	Generalizing Synecdoche
III	b ⊃ a	Yba * Aab	YbAa	Narrower Term some-all rel.	hyperonymous	Particularizing Synecdoche
IV	b X a	Yba * Yab	YbYa	Almost Generic some-some rel.	analogous	Metaphor (II + III)
V	b H a	Eba * Eab	EbEa	(related only in anthonym meaning)	outside	Lithote

Figure 3. Semantic comparisons for I-V relations

A more synthetic way to express such conjunctions is that of quantification of the predicate, as in column d of figure 3. To “quantify the predicate” we use the quantifier of the second categorical, which we have defined in terms of the inverted pair. E.g. case III can be read “only some b are all a”.

This double quantification have a contemporary application in Librarian Science, where the relation between index-words are ruled by “All-” and “Some-” quantifiers (for the past, see [10]). In cataloguing of subjects in Librarianship we find relations with the same logic structure of the 5 cases, as indicated in column e of figure 3.

In the linguistic-semantic fields as well as in the dictionaries of synonyms and contraries, the terms synonymous, hyponym, and hyperonym are known and used. We complete this casuistry by terms as “analogous” and “outside”, because less ambiguous than others like “meronym” or “antonym”.

In Rhetoric/Poetry (Liège's Group, T. Todorov), we can assimilate some “tropes” to the 5 cases (see [8]), as it can be seen in column g of figure 3.

The completion of the casuistry introduced by combination of distinctive predication allows the elimination of certain ambiguities of ordinary language, where the copula "is / are" makes no distinction between inclusion and identity, and of the same modern predicate logic, where the existential quantifier is chosen as primitive. All this allows the synthetic but also the exact expression of concepts in a wide spectrum of disciplines, and simplifies their understanding and ease of learning. The analysis of the concepts in distinctive terms brings out the ambiguities and helps the different cognitive background to confront.

3 THE 7 RELATIONS AND THE “DISTINCTIVE” COMPLEX SYLLOGISM

If the simultaneous comparison is not only of two sets but also of their two complementary sets we have to consider the context of the Universe of Discourse. There exists, indeed, a more complete way of dealing with concepts, whereby the negation plays a role, as in *not-b*, shown in grey in figure 4.

7 cases (b' is grey)	Distinctive Predicate Calculus: Iconic	Distinctive Predicate Calculus: compound predicates	Traditional Syllogistic: compound predicates	LINGUISTIC: SEMANTIC - DICTIONARY		LINGUISTIC: RETHORIC THROPHES	
1	bOa	Aba * Ab'a'	Aba * Ab'a'	synonymous		Definitio - Periphrasis	
2	b))a	Aba * Yb'a'	Aba * Ob'a'	hyponymous		Generalizing Synecdoche	
3	b((a	Yba * Ab'a'	Ab'a' * Oba	hyperonymous		Particularizing Synecdoche	
4	b)()(a	Yba * Yb'a'	Oba * Ob'a' * Iba * Ib'a'	tetrameronymous	analogous	Met. against a backdrop	Metaphor
5	b()a	Yba' * Ab'a	Eb'a' * Iba	hypercomplement		M. without a backdrop	
6	b)(a	Aba' * Yb'a	Eba * Ib'a'	hypocomplement	outside	contrary irony	Lithote
7	b$\wr$a	Aba' * Ab'a	Eba * Eb'a'	complement		oxymoron antithesis	

Figure 4. Semantic comparisons for 1-7 cases

The possible diagrammatic representations are then 7 in number (with the presupposition that b, a, not-b, not-a are different from the null class): see [2].

In the diagrams, the surrounding square represents the Universe of Discourse (UD). Only in case 5 this is missing, as the union of the two sets *a* and *b* form the UD.

We still have the possibility to characterize each case uniquely by conjoining two categoricals, but now we need a second conjunct that is a distinctive with the same subject and the same predicate as the first, but in the negative. A Boolean interpretation of the 7 cases is illustrated below:

1. $[(b \cap a) \neq \varnothing]*[(b \cap a') = \varnothing]*[(b' \cap a) = \varnothing]*[(b' \cap a') \neq \varnothing]$ that is $b = a$;
2. $[(b \cap a) \neq \varnothing]*[(b \cap a') = \varnothing]*[(b' \cap a) \neq \varnothing]*[(b' \cap a') \neq \varnothing]$ that is $b \subset a$;
3. $[(b \cap a) \neq \varnothing]*[(b \cap a') \neq \varnothing]*[(b' \cap a) = \varnothing]*[(b' \cap a') \neq \varnothing]$ that is $b \supset a$;
4. $[(b \cap a) \neq \varnothing]*[(b \cap a') \neq \varnothing]*[(b' \cap a) \neq \varnothing]*[(b' \cap a') \neq \varnothing]$;
5. $[(b \cap a) \neq \varnothing]*[(b \cap a') \neq \varnothing]*[(b' \cap a) \neq \varnothing]*[(b' \cap a') = \varnothing]$ that is $b \supset a'$;
6. $[(b \cap a) = \varnothing]*[(b \cap a') \neq \varnothing]*[(b' \cap a) \neq \varnothing]*[(b' \cap a') \neq \varnothing]$ that is $b \subset a'$;
7. $[(b \cap a) = \varnothing]*[(b \cap a') \neq \varnothing]*[(b' \cap a) \neq \varnothing]*[(b' \cap a') = \varnothing]$ that is $b = a'$

For all cases, the condition applies:

- $[b \neq \varnothing]*[b' \neq \varnothing]*[a \neq \varnothing]*[a' \neq \varnothing]$.

As regard the semantic plane, we may complete the splitting of “analogous” and “outside” cases, by adopting the relation of “complement” and by coining new expressions such as *tetrameronymous* for 4, or *hypercomplement* for 5 or *hypocomplement* for 6. If the five cases have introduced greater clarity in relations between concepts, 7 cases further refine this quality. For ex., the distinction between case 6 and case 7 is important, because represents the logical difference between contrariety (case 6), that is an incomplete incompatibility and admit intermediate or third objects, and contradiction (case 7), that is a complete incompatibility and represents a negation without intermediate or third objects.

If we use expressions taken from more ordinary language, a given term, with regard to a second, can be (from an extensional point of view and within a given UD): 1. an **Equivalent** of the second, 2. a **Restriction** of the second, 3. an **Expansion** of the second, 4. a **Limited Connection** of the second, 5. an **Integrative Connection** of the second, 6.a **Limited Disconnection** of the second, 7. an **Integrative Disconnection** of the second. The Complex Distinctive Calculus or Syllogism, D7c, which can be developed on this basis, is the object of another study by one of the present authors [4]. We show here (figure 5) a deductive table, where the second pair of terms (with inverted sign) is indicated by the double comma.

		1	2	3	4	5	6	7
		AbaA,,	AbaY,,	YbaA,,	YbaY,,	Yba'A,,	Aba' Y,,	Aba'A,,
1	AacA,,	AbcA,,	AbcY,,	YbcA,,	YbcY,,	Ybc'A,,	Abc' Y,,	Abc'A,,
2	AacY,,	AbcY,,	AbcY,,	*Ibc*	*Ycb*	Ybc'A,,	*Ib'c*	Ybc'A,,
3	YacA,,	YbcA,,	*Ib'c'*	YbcA,,	*Yc'b*	*Ibc'*	Abc' Y,,	Abc' Y,,
4	YacY,,	YbcY,,	*Yb'c*	*Ybc*		*Ybc*	*Yb'c*	YbcY,,
5	Yac'A,,	Ybc'A,,	*Ib'c*	Ybc'A,,	*Ycb*	*Ibc*	AbcY,,	AbcY,,
6	Aac' Y,,	Abc' Y,,	Abc' Y,,	*Ibc'*	*Yc'b*	YbcA,,	*Ib'c'*	YbcA,,
7	Aac'A,,	Abc'A,,	Abc' Y,,	Ybc'A,,	YbcY,,	YbcA,,	AbcY,,	AbcA,,

Figure 5. Deductive table of D7c

As a more iconic — that is, more visual and diagrammatic — alternative to the verbal-predicative system for the specific description of the seven cases, one may consider the Iconic Distinctive Calculus ID7 (see "iconic" column of figure 4, and [5,7]).

A numerical development of the iconic system is also possible (the subject is work in progress by the author, see [5, 6]). It can be considered the first step towards a Fuzzy Logic, in the sense of Zadeh [12].

As it has been shown elsewhere [4], it is possible to interpret the distinctive systems metalogically in terms of non-standard logics. For example, a sentence like "Only some x is P" (where some x stands for a set individuals and P for a predicate) can also be expressed as It is only partially true that the x's are P.

The most significant limitation of "syllogistic" deduction tied to the mono-argumental structure of its predicates (or bi-argumental, if seen as logic of classes). "Syllogistic" dictionaries cannot logically derive the relations of reciprocity, such as between father-son, husband-wife, etc. In perspective it is possible to build a multi-argumental distinctive logic, but at the moment it is outside our studies.

4 SEMANTIC APPLICATIONS TO LINGUISTIC TECHNOLOGIES

A central issue in the communication between two or more entities endowed with language (whether natural or artificial beings) is sharing the same vocabulary, whose terms / concepts are structured in a consistent way by logic point of view. The cognitive equipment or the same social behaviour, natural or simulated, in order to be implemented, requires a coherent conceptual apparatus, even if implicit.

The Distinctive systems combine exactness with flexibility, its closeness to natural language and the intuitiveness of its iconic-diagrammatical structure.

It could apply to dictionaries of synonyms, contraries, with the addition of intermediates, as well as to machine translation programs, so as to counter cognitive anomalies in the translations.

The adoption of the Distinctive Logic will improve the quality of all translations in a wide sense, not only of texts in the general usage of a natural language but also for specialist jargons, interdisciplinary codes, glossaries of technical manuals, thesauri, library indexes of all kinds. This is particularly timely in view of multiplication of hyper-specialist languages and terminologies of sub-disciplines. Scientific institutes feel the need of producing reliable translations for synthesis or interaction.

It will be helpful for organization of knowledge in the humanistic disciplines (philosophy, law, history, linguistics, etc.) and in scientific ones (mathematics, engineering, medical sciences, psychology, computer science and so forth). See [5,7].

5 ONTOLOGIES AND SEMANTIC SEARCH ENGINES

A Dictionary based on a Distinctive Syllogistic is an organic network of semantic relations between words (that can be representations of real objects or social entities), that can be seen as an Ontology. The premise of this construction deals with the Concepts / Words seen as sets or classes in terms of the five (or seven) distinctive relations. The elements of these sets are the features or characteristics of the Concepts / Words. We called "Semantic Prompter" this Ontology, because when it works as a filter for a traditional Syntactic Search Engine, together with the latter (and with a human input), it gives rise to a Semantic Search Engine. This operates on the basis of semantic similarities of the terms involved rather than on the basis of formal (phonological or spelling) similarities. For example, the word form *show* is more similar to the form *snow* than to its synonym *exhibit*.

The common dictionaries of Synonymy have a logical structure in too vague a sense: if *cat* is synonymous with *feline*, and even a *lion* is called *feline*, are *cat* and lion *synonyms*?

Some Ontologies like WordNet have a big limitation on the ground of the relations, observable on two levels: A) the logical relations between any pair of terms are just attributable to the simple or double inclusion/implication: there are no relations like our "Connection" or "Disconnection" (limitative or integrative). If they appear, they don't interact with the other relations, that is to say: do not constitute a deductive system; B) the totality of nodes or vertexes (= word - term - set) of the structure constitutes a hierarchical scheme, each term being such a parent (or predecessor) – son (or successor): the topological structure is a tree graph, there is no circuit, no cyclical arc and no network structure. If there are more trees, they aren't connected each other. See [1,11].

6 THE SEMANTIC PROMPTER ENGINE

In informatics terms the Prompter is a database/ program, a well-structured archiving system of notions, terms and data, with precise method and a specific question-answer algorithm. Querying the system, one will receive suggestions for further searches to be forwarded to the traditional engines.

When a normal search engine delivers unsatisfying or dubious results on a given word, the Prompter can be queried on this term. It can be queried, for example, with regard to the restrictions, limited connections, etc., or with regard to possible other terms validating totally, partially or not at all, the relation with the given term, within the terms of a given UD. From a logical point of view, such queries can correspond to the multiple request for the truth value resulting from the application of all possible categorical predicates to the term in question compared with other terms in the database, in combination with their negative counterparts. The subject class is definite — that is, without a quantifier — but the predication and the resulting truth value, as shown above, are definable on the grounds of the mere comparison of the features (attributes) of the two terms in question, whereby it is specified, if necessary, in terms of what UD the query is processed, with or without its complement. This comparison is carried out by means of the Boolean operators over feature sets. Once the answer list has been obtained, one will proceed to the selection of the new input term for the same search engine. The answers provide the *alternative terms* for new searches ordered by type (and relative grade) of similarity (or difference) according to the seven typologies (weak similarity in 4 or 5 cases, and no similarity for 6 or 7 ones). So we generate the *intermediate values* with regard to pure but flexible synonyms typical of human cognition and human language.

The user will thus not have to memorize or call up all correlated terms, at the risk of forgetting important ones, but will make his or her choice among the most promising options suggested, perhaps, if one wishes, till all options are exhausted. The combined action of the Prompter and the traditional search engine amounts to that of a *semantic search engine* with minimal human intervention. It represents both a limit and the opportunity to direct the search, as well as the possibility to discover unforeseen affinities or new associations based on objective common characteristics of concepts mostly thought to be far removed from each other.

To avoid the risk of lists becoming too long, answers will be organized hierarchically according to the level of generality desired or specified in the query. If so wished, the answers can be indexed according to level (0 horse, 1 equine, 2 mammal …). A further option, meant to simplify the selection of alternative terms, may consist in combining the output with a database that re-orders the output terms on the basis of frequency coefficients in the language or texts concerned.

6 DATABASE IMPLEMENTATION

The method for the construction of such a semantic network requires a category name for each node. There are two phases, both updatable: a) the tree, b) the creepers.

Phase a) build a tree structure with logical relations of inclusion/equivalence, and is useful in the early stages of cataloguing, structuring data and acquisition of concepts. Transitivity of inclusion and equivalence allow to automatically inferring any hierarchical knowledge which weren't explicit during the insertion.

Phase b) is necessary when we need to enrich the tree of new data it may be necessary a restructuring-extension of the model. The tree becomes a network through the introduction of two new logical relations: Connection and Disconnection. As the creepers attach themselves to branches that may be quite distant from each other, we have the possibility to attach the same feature to leaves and branching nodes on different branches. Thus Dolphin, which, we assume, will be on the Vertebrate branch along with, say, Lizard, will share the feature Marine with Starfish, which will be located on the Invertebrate branch.

The database will thus expand and update itself indefinitely, maintaining applicability in each phase. This allows it to integrate with other disciplines structured in analogous ways in other databases. The possibility of integrating a variety of disciplines will over time enhance the usage scope of the Prompter, turning it into a continuously updatable and improvable encyclopaedia.

The data source can be found, for ex. in the abstracts of scientific / literary articles, in the key-word or ISBN index catalogues. One of the authors, Daniele Ingrassia, has actually prepared a software that can extract meta-data from Wikipedia and organize them by synonymy, hyponymy, hyperonymy on the basis of proximity and statistical recurrences.

REFERENCES

[1] Alesso P. H. and Smith C. F.: Thinking on the Web: Berners-Lee, Gödel and Turing, Winley & Son (2006)

[2] Bird O.: Syllogistic and its Extensions. Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1964)

[3] Blanché R.: Structures Intellectuelles: Essai sur l'Organisation Systematique des Concepts. J.Vrin, Paris (1966)

[4] Cavaliere F.: Fuzzy Syllogisms, Numerical Square, Triangle of Contraries, Interbivalence. In J.Y. Beziau and D. Jacquette (eds), Around and Beyond the Square of Opposition, Springer, Basel, 241-260 (2012)

[5] Cavaliere F.: Suggeritore Semantico per Motori di Ricerca e Traduttori, La Feltrinelli (2013 a)

[6] Cavaliere F.: A Diagrammatic Bridge between Standard and Non-Standard Logics: the Numerical Segment. In: Visual Reasoning with Diagrams, Springer, Basel, 73-81 (2013 b).

[7] Cavaliere F.: . A "Distinctive" Logic for Ontologies and Semantic Search Engines. In: Representation and Reality: Humans, Animals and Machines (to be appear).

[8] Cerisola P.L.: Trattato di retorica e semiotica letteraria. La Scuola, Brescia (1983)

[9] Faris J.A.: The Gergonne relations. The Journal of Symbolic Logic 20, 207-31 (1955)

[10] Hamilton W.: Lectures on Logic. Edinburg (1860)

[11] Kashyap V., Bussler C., Moran M.: The Semantic Web: Semantics for Data and Services on the Web (Data-Centric Systems and Applications), Springler (2008)

[12] Zadeh L. A.: A Computational Approach to Fuzzy Quantifiers in Natural Languages, Computer & Mathematics with Applications, 9 (1), 149-184 (1983).