Archeologia e Calcolatori
13, 2002, 217-232

ESTIMATING SUBJECTIVITY OF TYPOLOGISTS AND TYPOLOGICAL CLASSIFICATION WITH FUZZY LOGIC

1. Classification

A basic concern of prehistorians is the delineation of cultural configurations, or patterns, from assemblages (Trigger 1968, 15). In order to achieve this goal, artifacts are analyzed by defining a system of ordered units (as typological groups), used for a comparison between assemblages. The primary unit for comparison is the **component**, defined as a single occupation in the history of a given site (Willey, Phillips 1958, 21). For a component to be meaningful, the time segment that is isolated must delimit a static cultural situation, without any significant change. Thus, the cultural material associated with this component should represent the people who inhabited the site, at a given point in their cultural development.

Intuitively, a major component of prehistoric material culture, flint tools, are defined as: artifacts with intentional retouch (a series of removals), intended either for the preparation of a hafting or a working edge. The system of ordered units, our typology, follows the rules of taxonomy:

1. It is based on the identification of certain forms and of certain characteristics.
2. The chosen characteristics have to be exclusive, present at each case of type.
3. Each case can be classified in **one and only one** type.

Two essential properties are characteristic of types: identity, or identifiability, and a meaning, relevant to some purpose. One of the major "weaknesses" of typological research is the nature itself of types – being defined partly intuitively and partly rationally, partly essential and partly instrumental, most typological lists are polythetic, so there are no fixed criteria of "typehood".

Archaeological types are usually distinguished by norms or central tendencies, rather than distinct boundaries (Adams, Adams 1991, 20). The main essence of types is whether they serve a specific (scientific) purpose, or not. Every type necessarily has to have a diagnostic attribute, or a cluster of attributes, that sets it apart from all other types.

A valid classification approach is the one often called "lumping" (Dunnell 1971; Everitt 1974), in which the emphasis is on "external isolation" of types (Adams, Adams 1991), rather than searching for minor variations, which can be the result of "internal cohesion" among similar entities ("splitting"). A basic assumption is that although types are defined *a posteriori*, as a result of criteria used by typologists, they also existed *a priori*, since they can be identified and described.

Several approaches to typological studies were adopted since archaeology raised as a science (for a good summary, see WHITTAKER *et al.* 1998). Following is a survey of main ideas adopted through years in the typological research, and their application.

One approach to the concept of *fossile directeur* (supposedly defining components of a given culture, usually particular tool-types) was formulated by Gallus, who described (prehistoric) culture as «…an assemblage of objects and observable manifestations or traits…which according to form and expression (observable externalization) can be regarded as different (disparate) from any other such assemblages of traits…the objects and traits, which document a culture, are the synchronic end results…of a typological series in a particular time period…» (GALLUS 1977, 143). In this sense, the traditional cultural historic view of types, defining cultural *fossiles directeurs*, is adopted in its broader sense, including other aspects of lithic industries. Thus, types with potential cultural identifiability were used to classify assemblages in a time-scale of developing cultures.

The possibility of using typological studies to explore cognitive capabilities of past humans was expressed since the sixties in works of CHANG (1967, 228 – "cognitive system of the makers") or COWGILL (1977, 236 – "alien system of cognition"). These concepts were adopted by Gallus for his definition of an "organic typology", where the concept of a "tool-type", as formulated by the prehistory researcher in his mind, «…should match as closely as possible the idea formed by the prehistoric artisan himself…» (GALLUS 1977, 134). The "organic typology" is based therefore on a comparison (matching) between the ancient "mental template" (of the knapper) and the modern one (of the prehistorian). The use of typology is based on the assumption that human brain, when confronted with a variety of perceptions in his environment, notices repeated occurrences of identical or similar occurrences, forming cognitive units, or engrammes (a complex of neural networks, which reflect an idea, or a concept) (MARGENAU 1950, 58). Thus, the aim of typological analysis is to establish tool-types that reflect «…the concept of a prehistoric artisan, externalized in the form of an actual tool, in order to serve a **need**, to perform a **task**…» (GALLUS 1977, 134).

This view contrasts Tixier's, who declared that "tools" are artifacts that prehistorians define as such, and not what prehistoric people might have regarded as such (TIXIER 1967, 815). This view still dominates the scientific prehistoric research, even though it was criticized that «…it dehumanizes prehistory…we…are in danger of arriving at historically anachronistic results…» (GALLUS 1977, 135).

Already in the 60's, Leroi-Gourhan pointed to the problematic of the "classic" (of those days) typologically oriented prehistoric research, proposing a research framework, the *chaîne opératoire,* aimed to produce a "biology of techniques" – a structural and functional awareness of techniques; he formulated the "means of action on matter", which included terms as "percussion", "abrasion", etc. He concluded that «techniques are both gestures and tools, organized in a true syntax» (Leroi-Gourhan 1964, 164). The act of fabrication is a dialogue between the knapper and the worked material (Leroi-Gourhan 1965, 132) implicating the natural determination of the *tendance* and the cultural idiosyncrasy of the *fait* (Leroi-Gourhan 1943, 23-43).

Since the beginning of this century, and especially during the fifties, under the influence of François Bordes, typological research was regarded as the main "objective" tool in establishing chronological frameworks and comparisons between sites. This view is best reflected in Bordes's work: «…la typologie est la science qui permet reconnaître, de définir, et de classer les différentes variétés d'outils se rencontrant dans les gisements» (Bordes 1961). Selected flint tool types were regarded as *fossiles directeurs*, which could be used for dating archaeological horizons, in a way similar to fossils, used for dating geological layers.

The recognition of repetitive morphological forms enabled researchers to group them into "types", defined by their characteristic "retouch". A special vocabulary was needed for the definition of types, the terms being taken from ethnographic analogies, mode of preparation or supposed function (e.g. burin, scraper, etc.). For a more detailed classification of tools, sub-types were defined, according to specific morphological features (e.g. convergent scraper), modes of preparation (e.g. dihedral burin) or toponymical belonging to a particular site (e.g. Kebara point).

At this stage of research, most of the effort was directed towards the definition of chronological sequences, based on the *fossile directeur* concept that was supposed to define cultural entities. Following a rapid development in prehistoric research, a need for standardization of the terminological vocabulary was needed (Brezillon 1968). Several symposiums were held, for the definition of a universal "typological list", suitable for all regions, for a given specific time period (Hours 1974).

A further development of the prehistoric research followed the observations of Bordes, which concluded that *fossiles directeurs* are meaningless in their singular forms, and a method of percentage comparison between assemblages should be developed (Bordes 1950, 1961). The idea was to define a representative typological list, a complete inventory of tools characteristic of a given region/period and compare it (graphically) with inventories of other sites. Parallelisms and differences between the (cumulative) graphs were used for definition of cultures and cultural regions.

During this period, as a consequence of an increasing use of statistical methods, alternative typological approaches were proposed, as the analytical typology of Laplace (1966) or the descriptive morphology of Leroi-Gourhan (Leroi-Gourhan, Laming-Emperaire 1966). The main assumption underlying the Laplacian approach is that cultural changes can be explained by a process of "predetermination" – the analysis of flint tools should be based on an hierarchy of "complete characteristics", build on discrete and continuous "variables" (Laplace 1966). Thus, each item is not only nominated (scraper, burin, etc.), but also morphologically described, by a series of codes. Thus, the "subjective" aspect of the typological definition is eliminated, each item being classified using the mathematical aid of factorial analysis and automatic classification.

This somehow promising approach, to eliminate the "subjectivity" from typological studies, was forwarded by intensive researches in the application of various statistical methods of identifying, defining and classifying past objects into meaningful types. This research was oriented in several directions, including elaborate data-extraction through visualization in graphic displays (e.g. Hovers, Raveh 2000), evaluation, using Bayesian methods, of "subjectivity errors of typologists" (e.g. Whittaker *et al.* 1998; Read 1989; Gnaden, Holdaway 2000) or the application of various statistical methods to "eliminate" the subjective side of the type's nature (Bailey 1994; Read, Russell 1996; Shott 2000, Djindjian 2001).

Despite these efforts (see also Bisson 2000), it seems that any tentative to separate between "subjectivity" and "typology" may have limited results, given the nature itself of typological research: «…le bon typologiste perçoit avec l'expérience après avoir analysé quelques milliers de pièces…», spending a good deal of time «…au course de longs tête-à-tête avec les outillages lithiques…» (Demars, Laurent 1992, 20).

2. The subjective nature of "objective" classification

Apparently, a part of the classification process is an inherently subjective one, intuition having a larger or lesser influence, which depends on the researcher and on the classification method (the typological list) applied. Thus, "objective" methods are largely subjective, consisting upon the measurement of certain parameters, often arbitrarily defined, allowing in many cases a great degree of freedom, having a loose definition. For example, given some characteristic C, a classification based on C requires measuring a set of physical parameters $p_1, p_2, \ldots p_n$ and verifying if

$$f_C(p_1, p_2, \ldots p_n) = 1 \text{ or not}$$

f_C being some predefined function that completely describes the characteristic C.

In most cases, the definition of the classification function will be very simple, for instance it may require simply that p_i lies in a given interval P_i, so that in this case

$$f_C(p_1, p_2, \ldots p_n) = \sum_k I_k^2(p_k)$$

I_k being the characteristic function of the interval P_i, that is

$$I_k = \begin{cases} 1 & \text{for } p_k \in P_k \\ 0 & \text{for } p_k \notin P_k \end{cases}$$

It is simple to verify that any "objective" definition of tools may be easily translated in this way: for instance, the definition of "scraper" (the characteristic C of our case), according to the Laplace method, is an item with a continuous, simple or scaled retouch. Thus, two parameters defines C:

I. p_1 = delineation of retouch, with the value of "continuous", described as a retouch that draws a continuous line, straight, concave, convex or sinuous along the edge of the item;

II. p_2 = mode of retouch, with two potential values: "simple" (along the edge of a thin flake, maintains the cutting edge, the retouch being a series of more or less elongated scars, sometimes *en écaille*, which form with the striking plane of the item a narrow angle); "surélevée" (along the edge and the face of a thick flake, conserving or not the cutting edge, the retouch being irregular, wide and scalariforme, with sub-parallel bladelet scars).

Therefore C, the scraper in our example, is a tool on a thin or thick flake, with a retouch varying from simple to scalariforme, covering some or all the cutting edge, modifying it or not. The ambiguity of this definition is self-evident, especially when confronted with definitions of other tool-types, as retouched flakes or truncations (see also Tixier 1967).

This process gives a complete illusion of "objectivity" being based, as it is, on the absoluteness of mathematical formulae, and as far as mathematics is applicable, it is so. However, the application of such a system requires some operations that are outside the realm of mathematics, and here the objectivity goes awry.

First of all, who decided that f_C characterizes the class C? Did prehistoric industries have a quality control that discarded tools not obeying to this production function, which was thereafter communicated to modern archaeologists so that they could check the product characteristics and infer the product name, type and purpose? Often, as noticed before, it is a modern statistical analysis that groups together different artifacts recognizing (possibly after having applied statistical analysis) common characters that may also correspond to some hypothized "use" rule (e. g. from wear) and abstracting these results to obtain a general rule, which is then applied to other assemblages. Further deductions are then inferred by this classification and the process continues, basing repeatedly on simplifications, approximations, errors and inference. So the process that leads from artifacts to theories does not only rely on numbers and computations, but has to cope with decisions: if the prescribed length for a class is 4 cm, should we accept something 4.1 cm long? And something 4.2 cm long? And when should we stop accepting and start rejecting?

The fact is that items to be classified may be grouped in three heaps: what certainly does belong to that class, what for sure does not and what is uncertain.

The only validation that can be applied is the acceptance by the scientific community: every scholar may, in principle, check the deductions of another researcher, both from what he or she communicated and – even if with greater difficulty – on the "real thing", the objects that were accurately collected and documented; verify by subjective judgment if the results are convincing; argue (publicly or privately) in favor or against; then the conclusions are accepted or rejected. In this chain, statistical and mathematical considerations have to be convincing, not decisive, even if they play a substantial role in deciding, classifying and concluding. So, any contribution to the transparency of computations supports the scientific correctness of the process.

What usually happens is that "uncertain" objects are assigned to some class and then deductions are made basing on the count, percentage or some other global index of the number of objects belonging to different classes. Thus, wrong assignments may lead to erroneous results. Still worst conditions may derive by the use of a computer, which leaves no space to uncertainty and hence forces the researcher to assign objects to one determinate class, by ticking some box and inserting the value into a database: with the automatism of its deductions and the impossibility of computational errors, a computer adds to the false sense of reliability of the final result.

The above considerations do not mean that all the theory of classification based on statistical analysis is to be rejected, but only that extreme caution must be exerted when basing deductions only on statistical computations; they moreover suggest that the method that will be introduced later may help to make visible the uncertainty that underlies any such deduction and help to evaluate some numerical reliability index.

3. Fuzzy set theory

Fuzzy sets, first introduced by Zadeh in 1965 (Zadeh 1965), generalize the familiar concept of sets extending the indicator function. As stated above, the indicator function of a set is a function that values 1 within the set, 0 outside. Since a set *is* its indicator function, a fuzzy set is defined as a function with values in the interval [0, 1], namely given some (traditional) set *U*, a fuzzy subset of *U* is a function

$$f : U \mapsto [0,1]$$

If the only possible values are 0 and 1, this reduces to the classical definition of a subset.

As for classical set theory, this traduces into logic, the truth function of a predicate being the fuzzy set corresponding to its truth values.

Fuzzy set theory has been thoroughly studied (see, for instance, Yager, Filev 1994 and its extensive bibliography, and Li, Yen 1995) and has many applications in engineering. Usually, the goal is the defuzzification, that is the reduction of control problems to the non-fuzzy case. Our case is exactly the opposite.

Fuzzy logic has seen also many applications in linguistic and other sciences, being able to manage situations in which binary logic is inapplicable.

As far as archaeology applications are concerned, after a first paper by Barceló (1996) fuzzy set theory and fuzzy logic has been applied to databases, creating and implementing a RDBMS model with fuzzy functions and using it to study an Etruscan cemetery, and to GIS: spatial relations are a good example of fuzzy statements as in the statement "this place is near that one" (for both applications, see Niccolucci *et al.* 2001 and Crescioli *et al.* 2000 and the references quoted there).

For the present problem we can use the fuzzy set model defining the grade of belonging of the item *i* to class A as a number a_i, with $0 \leq a_i \leq 1$, which expresses the subjective degree of belief that the assertion "item *i* is a member of class A" is true.

This has some kinship with probability theory, but is not identical to it: the main difference is that no normalization condition of the type $\Sigma\, a_i = 1$ is imposed. The numbers a_i are merely a way of communicating how reliable the assignment is from the researchers' viewpoint. They help them to do a better overall analysis and help other researchers to evaluate the authors' results. As it will be seen in what follows, the use of fuzzy set theory may also solve conflicts of attribution, that is situations in which different researchers classify differently the same assemblage of objects, resolving in different ways the classification of uncertain artifacts.

4. Using fuzzy set theory for the classification of artifacts

The fuzzy method may be easily applied to flint tool classification as a generalization of the usual methodology. So far, the result of classification consists of a list of numbered items (the finds) accompanied by a label (the tool class assigned to it). Then statistical computations may be applied, as simple counts or percentage evaluation, clustering or whatever the researcher estimates as more significant for the investigation to be carried on.

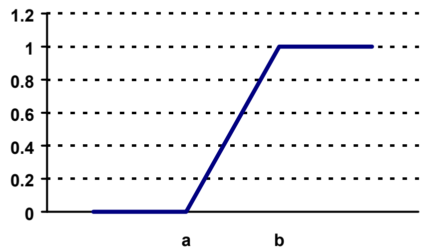
With the fuzzy method, the starting point of this statistical processing is, instead of a one-dimensional list, a matrix *A* having all the items as rows and all possible tool classes as columns. Each element a_{ij} of the matrix is a number in [0, 1] expressing the degree of possibility the evaluator assigns to item *i* belonging to tool type *j*. If for each row there is only one $a_{ij} = 1$ and all the others are 0, this method reduces to the classical one.

Table 1 gives an example of fuzzy classification for an assemblage of 50 flint tools. Starting from this table, various statistical analysis may be performed; for instance the relative presence of different types requires computing the column totals and evaluating the relative weight of each type as a fraction or percentage of the total, as shown in the last line of the table.

To evaluate the fuzzy coefficients, the researcher has several alternatives. Some of these methods are outlined in Li, Yen (1995) as far as subjective evaluation is concerned; when measurements are involved, a reliability function as the following may be used. Let us suppose that some type assignment is based on a numeric parameter *x*, such that the item belongs to the type under consideration if $x > b$, b being a numeric threshold established in previous research. Let us also assume that it is unquestioned that if $x < a$ the item does not belong to this type (if no such value exists, we may assume $a = 0$). A simple reliability function *f* may then be created as follows:

$$f(x) = \begin{cases} 0 & \text{for } x < a \\ \dfrac{x-a}{b-a} & \text{for } a \le x \le b \\ 1 & \text{for } x > b \end{cases}$$

This is equivalent to use a linear function as shown by the following graph.



Of course the latter method may result rather cumbersome to use manually, but it may be easily implemented when using a computer. This may occur when finds are recorded automatically or semi-automatically using, for instance, a 3D-scanner, and measurements are done by the software on the resulting digital 3D model. The method may be adapted to take into account systematic errors of the automatic method, due to inaccuracy in scanning, round-off, and so on.

The above method may be used also when more than one numeric parameter is involved. If, for instance, two measurable features have to be used, after evaluating each reliability coefficient f_1 and f_2, to compute the global reliability coefficient f one may adopt a "pessimistic" viewpoint

$$f = \min(f_1, f_2)$$

(the global reliability is the worst of the two), which is equivalent to require an AND condition between the fulfillment of the two requirements; or an "optimistic" viewpoint

$$f = \max(f_1, f_2)$$

(the global reliability is the best of the two), which corresponds to an OR condition between the two distinct requirements in the Boolean traditional framework. Also intermediate values have been proposed and they may be used to give a greater weight to the principal parameter and a lesser importance to a secondary one.

In the experiment shown in the following paragraph, we adopted a much simpler method: ask the expert who is classifying the artifacts to express a subjective "mark" for each type, basing on his or her experience. Subjective as it may be, this method generalizes what is done in most cases: evaluate tools by observation and experience. It simply adds to common practice the requirement of numerically expressing the referee's confidence in the typology just assigned and to list other possible types, each one with its own confidence index.

5. The case study

To verify the impact of the fuzzy classification method we chose an assemblage of 50 flint tools from a proto-historic site in Southern Israel. The material originates from a single occupation phase; therefore intrusions were basically absent.

This set of artifacts was given to 5 different experts and they were invited to classify them according to 12 classes and a residual ("Varia") one. The result is shown in Table 1. As it can be seen, the agreement on classification was not complete even if the researchers came from the same "school", had similar training and experience and had a good expertise on the typology from which the sample derived. Of course it was assumed that no mistake was done.

This posed a series of problems: who should be the "trusted" expert? How could uncertainty be managed? And then, how scientifically reliable

	Referee 1	2	3	4	5
Item 1	Denticulated	Scraper	Scraper	Denticulated	Denticulated
2	Burin	Burin	Scraper	Truncation	Burin
3	Retouched Flake	Retouched Flake	Borer	Borer	Borer
4	Notch	Notch	Retouched Flake	Notch	Notch
5	Retouched Flake	Denticulated	Sickle Blade	Tabular Scraper	Denticulated
6	Denticulated	Denticulated	Denticulated	Borer	Borer
7	Notch	Denticulated	Scraper	Notch	Denticulated
8	Scraper	Notch	Notch	Denticulated	Notch
9	Retouched Blade	Retouched Flake	Scraper	Retouched Flake	Retouched Flake
10	Scraper	Retouched Blade	Sickle Blade	Denticulated	Retouched Blade
11	Borer	Denticulated	Denticulated	Denticulated	Denticulated
12	Borer	Scraper	Scraper	Retouched Flake	Denticulated
13	Scraper	Scraper	Retouched Flake	Retouched Flake	Retouched Flake
14	Retouched Flake	Retouched Flake	Truncation	Retouched Flake	Retouched Flake
15	Borer	Retouched Flake	Borer	Notch	Notch
16	Scraper	Scraper	Truncation	Notch	Denticulated
17	Truncation	Retouched Flake	Tabular Scraper	Retouched Blade	Retouched Flake
18	Notch	Scraper	Scraper	Notch	Retouched Flake
19	Retouched Blade	Retouched Blade	Sickle Blade	Sickle Blade	Retouched Blade
20	Retouched Flake	Borer	Borer	Retouched Flake	Notch
21	Scraper	Scraper	Retouched Flake	Notch	Scraper
22	Sickle Blade	Sickle Blade	Sickle Blade	Scraper	Sickle Blade
23	Bifacial	Bifacial	Denticulated	Denticulated	Varia
24	Borer	Borer	Borer	Borer	Borer
25	Retouched Blade	Sickle Blade	Sickle Blade	Retouched Blade	Sickle Blade
26	Retouched Blade	Sickle Blade	Sickle Blade	Varia	Retouched Blade
27	Retouched Blade	Retouched Blade	Scraper	Notch	Retouched Blade
28	Notch	Notch	Borer	Notch	Notch
29	Notch	Scraper	Denticulated	Notch	Scraper
30	Retouched Flake	Retouched Blade	Scraper	Varia	Retouched Flake
31	Scraper	Scraper	Retouched Flake	Scraper	Scraper
32	Scraper	Scraper	Scraper	Retouched Blade	Varia
33	Notch	Retouched Blade	Scraper	Notch	Retouched Blade
34	Retouched Flake	Burin	Sickle Blade	Retouched Flake	Varia
35	Borer	Scraper	Borer	Borer	Borer
36	Notch	Retouched Flake	Scraper	Truncation	Retouched Flake
37	Notch	Scraper	Denticulated	Retouched Flake	Retouched Flake
38	Scraper	Denticulated	Denticulated	Denticulated	Denticulated
39	Notch	Scraper	Borer	Denticulated	Notch
40	Notch	Retouched Blade	Scraper	Notch	Retouched Flake
41	Retouched Bladelet	Retouched Bladelet	Sickle Blade	Varia	Retouched Blade
42	Truncation	Retouched Flake	Truncation	Retouched Flake	Retouched Flake
43	Scraper	Scraper	Scraper	Scraper	Denticulated
44	Truncation	Denticulated	Denticulated	Denticulated	Denticulated
45	Borer	Borer	Borer	Borer	Borer
46	Retouched Flake	Scraper	Scraper	Denticulated	Denticulated
47	Truncation	Scraper	Borer	Retouched Flake	Retouched Flake
48	Truncation	Retouched Flake	Truncation	Truncation	Retouched Flake
49	Borer	Retouched Flake	Borer	Borer	Varia
50	Scraper	Scraper	Denticulated	Denticulated	Denticulated

Table 1 – Results of typological classification of an assemblage containing fifty tools, by five different researchers. Note the difference among referees' inventory lists.

Item No.	Scraper	Tabular Scraper	Borer	Burin	Truncation	Retouched Flake	Notch	Denticulated	Retouched Blade	Sickle Blade	Retouched Bladelet	Bifacial	Varia	***R***
1	0.7							0.9						**0.51**
2				1.0										**1.00**
3						1.0	1.0	1.0						**0.33**
4							1.0							**1.00**
5						0.8	0.8	0.5						**0.30**
6			0.5					1.0						**0.67**
7						0.8	1.0							**0.56**
8	0.8							0.8				0.5		**0.30**
9									1.0	0.5				**0.67**
10	0.7		0.5					0.7						**0.26**
11			0.9			0.5		0.5						**0.43**
12			0.9			0.5		0.5						**0.43**
13	1.0													**1.00**
14						1.0								**1.00**
15			0.8											**0.80**
16	0.8				0.5									**0.49**
17					1.0									**1.00**
18							1.0							**1.00**
19									1.0	1.0				**0.50**
20			0.5			1.0								**0.67**
21	0.8				0.5									**0.49**
22										1.0				**1.00**
23												1.0		**1.00**
24			0.8				0.5							**0.49**
25									1.0	1.0				**0.50**
26									1.0	0.5				**0.67**
27									1.0					**1.00**
28							1.0							**1.00**
29							1.0					0.5		**0.67**
30						0.8								**0.80**
31	1.0													**1.00**
32	0.7					0.5								**0.41**
33							1.0		1.0					**0.50**
34						1.0								**1.00**
35			1.0											**1.00**
36							0.7							**0.70**
37							1.0							**1.00**
38	1.0							1.0						**0.50**
39			0.5				0.7							**0.41**
40					0.5		1.0							**0.67**
41											1.0			**1.00**
42					0.8		0.5							**0.49**
43	1.0							0.5						**0.67**
44					0.8			0.5						**0.49**
45			1.0											**1.00**
46						1.0		1.0						**0.50**
47	0.5				1.0									**0.67**
48					1.0									**1.00**
49			1.0											**1.00**
50	1.0							1.0						**0.50**
Total	10.0	0.0	8.4	1.0	6.1	8.9	12.2	9.9	6.0	4.0	1.0	2.0	0.0	
%	14.4%	0.0%	12.1%	1.4%	8.8%	12.8%	17.6%	14.2%	8.6%	5.8%	1.4%	2.9%	0.0%	

Table 2 – Example of fuzzy assignments by referee 1 and corresponding values of *R*. The average value of *R* is 0.70.

Referee	Scraper	Tabular Scraper	Borer	Burin	Truncation	Retouched Flake	Notch	Denticulated	Retouched Blade	Sickle Blade	Retouched Bladelet	Bifacial	Varia
Traditional (non-fuzzy) method													
1	**20.0%**	0.0%	14.0%	2.0%	10.0%	14.0%	**20.0%**	4.0%	10.0%	2.0%	2.0%	2.0%	0.0%
2	32.0%	0.0%	6.0%	4.0%	0.0%	18.0%	6.0%	12.0%	12.0%	6.0%	2.0%	2.0%	0.0%
3	28.0%	2.0%	20.0%	0.0%	8.0%	8.0%	2.0%	16.0%	0.0%	16.0%	0.0%	0.0%	0.0%
4	6.0%	2.0%	12.0%	0.0%	6.0%	18.0%	22.0%	20.0%	6.0%	2.0%	0.0%	0.0%	6.0%
5	6.0%	0.0%	10.0%	2.0%	0.0%	**24.0%**	12.0%	**22.0%**	12.0%	4.0%	0.0%	0.0%	8.0%
Average	18.4%	0.8%	12.4%	1.6%	4.8%	16.4%	12.4%	14.8%	8.0%	6.0%	0.8%	0.8%	2.8%
Fuzzy method													
1	**14.4%**	0.0%	12.1%	1.4%	8.8%	12.8%	**17.6%**	14.2%	8.6%	5.8%	1.4%	2.9%	0.0%
2	25.7%	0.0%	6.9%	3.7%	0.0%	17.1%	12.1%	14.1%	9.9%	5.4%	1.4%	3.8%	0.0%
3	23.8%	1.7%	20.0%	1.2%	5.4%	14.6%	5.2%	12.3%	2.8%	10.9%	0.0%	2.0%	0.0%
4	4.7%	1.9%	12.9%	0.8%	7.9%	14.4%	23.9%	19.5%	4.7%	3.8%	0.0%	0.0%	5.6%
5	5.6%	0.0%	7.3%	1.7%	0.0%	**22.5%**	13.6%	**27.0%**	9.3%	3.3%	0.0%	0.7%	9.1%
Average	14.9%	0.7%	11.8%	1.7%	4.4%	16.3%	14.5%	17.4%	7.1%	5.8%	0.6%	1.9%	2.9%

Table 3 – Percentages of assignments to types by referee, with the traditional method (% of cases in which a typology is the chosen one) and with fuzzy method (normalized value to 100 of the sum of fuzzy coefficients, shown in the last line of Table 2). Boldface evidences cases in which a second choice becomes the first using the fuzzy method.

could classification be considered, since it was not replicable even in so simple "laboratory" conditions?

The fact is that as previously discussed classification does not lead to a precise, "crispy" result, but carries some amount of uncertainty that cannot be completely eliminated. In most cases, this has no consequence on the final result, but sometimes it may make the difference. Thus, the fuzzy method is proposed as a possible solution.

The experts were asked to extend their classification expressing any "possible" typology assignment by means of a numeric coefficient in the range [0, 1], 0 meaning that they did not consider such an assignment as possible, 1 meaning that they considered the related typology as most reliable. No constraint was set on the number of 1's, so that multiple "most reliable" assignment were allowed: this is no contradiction, since it may happen that after excluding a number of classes it is almost non-decidable which one of the remaining must be chosen, all appearing as equally reasonable.

This led to the compilation of 5 tables, one for each referee, showing different assignments accompanied by the reliability coefficients assigned by each expert, which could then be used to obtain a joint classification table showing the compound reliability coefficients. For example, Table 2 reports the values assigned by referee 1.

The final assignments and coefficients may be calculated by means of a weighted average of each evaluator's coefficients. This was done with equal weights, obtaining the values shown in Table 3.

6. The impact of the fuzzy method

It might be argued that the differences between the usual classification method and the fuzzy one are not so relevant, and that they do not justify the waste of time.

The first reason of using the fuzzy method is to make comparable different assignments, as came out by our experiment, without having to choose between different referees.

Second, there are differences, which make one typology prevail in the modal ranking only if the fuzzy method is used. This happens if there is a typology that is often the second choice in a referee's classification, while the first choice is always a different one: if this is taken into account, as the fuzzy method allows, the marks for the second choice add up obtaining a total score which makes such a typology relevant in the global analysis. This result is possible only with the fuzzy method, since with the usual one second choices are always discarded and only the first one is considered. (See table 3).

The fuzzy classification method helps also to identify "difficult" items. For this we can define a reliability index of each item which takes jointly into account the spread of possibilities assigned to that item (i.e. the number of different typology choices considered as possible for it) and the value of the most reliable assignment.

The value we propose for this index is given by the following formula:

$$R(x_1, x_2, \ldots x_m) = \frac{\left(\max_k x_k\right)^2}{\sum_k x_k}$$

relating the reliability index R of a set of fuzzy coefficients x_1, x_2, $\ldots x_m$ to the maximum value of the x's and inversely to their sum. R is defined as zero if all the x_k are zero. The reliability index is a number in the interval [0, 1] with some interesting properties that make it a good candidate for this role.

R is 1 only if the set of x's is a crispy set, namely if the assignment has no fuzziness: one and only one of the x is 1, all the others are zero. As a function of the coefficients, R is a continuous function, in particular when all x approach zero. Moreover, R decreases if the number of non-zero fuzzy coefficients increases, thus increasing the fuzziness of the classification; R increases, as expected, if the maximum value of the fuzzy coefficients is greater, reducing the fuzziness by expressing more confidence in the preferred typology. Thus R may be used to evaluate the global reliability of type assignments for an item and hence the "difficulty" of that specific item to be classified. A global R for an assemblage can be easily defined by averaging R over the items: the result is an average index of the reliability of the assemblage as a whole.

7. Conclusions and further work

It has been shown that applying concepts of fuzzy logic to typological classification of artifacts may aid in various aspects to a better interpretation of past relics. Future research will focus on two main aspects: evaluation of *R* coefficient of "fuzzy" types in main periods/cultures and subsequently, evaluation of *R* at assemblage level, identifying characteristic *R*'s for various assemblages representing different cultures and periods. Moreover, it is intended to explore the fuzziness not only of tool assemblages, but for waste products as well, hoping to achieve a simple level of elaboration of *R* in a way that synthesis of research and accuracy of explanations will be easily achieved.

Sorin Hermon
The Archaeology Division
Ben-Gurion University of the Negev

Franco Niccolucci
Dipartimento di Urbanistica e
Pianificazione del Territorio
Università di Firenze

REFERENCES

Adams W.E., Adams E.W. 1991, *Archaeological Typology and Practical Typology*, Cambridge, Cambridge University Press.

Bailey K.D. 1994, *Typologies and Taxonomies*, Thousand Oaks, CA., Sage University Press.

Barceló J.A. 1996, *Heuristic classification and fuzzy sets. New tools for archaeological typologies*, in H. Kamermans, K. Fennema (eds.), *Interfacing the Past*, Analecta Praehistorica Leidensia 28, 155-164.

Bisson M.S. 2000, *Nineteenth century tools for twenty-first century archaeology? Why the Middle Paleolithic typology of François Bordes must be replaced*, «Journal of Archaological Method and Theory», 7 (1), 1-48.

Bordes F. 1950, *Principe d'une méthode d'étude des techniques et de la typologie du Paléolithique ancien et moyen*, «L'Anthropologie », 54, 19-34.

Bordes F. 1961, *Typologie du Paléolithique Ancient et Moyen*, Bordeaux, Delmas.

Brezillon N.M. 1968, *La dénomination des objets de pierre taillée*, Paris, CNRS.

Chang K.C. 1967, *Major aspects of the interrelationship of archaeology and ethnology*, «Current Anthropology», 8, 227-243.

Cowgill G.L. 1977, *The trouble with significance tests and what can we do about it*, «American Antiquity», 42, 350-368.

Crescioli M., D'Andrea A., Niccolucci F. 2000, *A GIS-based analysis of the Etruscan cemetery of Pontecagnano using fuzzy logic*, in G.R. Lock (ed.), *Beyond the Map: Archaeology and Spatial Technologies*, Amsterdam, IOS Press, 157-179.

Demars P.Y., Laurent P. 1992, *Types d'outils lithiques du Paléolithique Supérieur en Europe*, Paris, CNRS.

Djindjian F. 2001, *Artefact Analysis*, in Z. Stanèiè, T. Veljanovski (eds.), *Computing Archaeology for Understanding the Past*, Oxford, Archaeopress, 931, 41-52.

Dunnell R.C. 1971, *Systematics in Prehistory*, New York, Free Press.

EVERITT B. 1974, *Cluster Analysis*, New York, John Wiley & Sons.

GALLUS A. 1977, *Organic typology*, in R.V.S. WRIGHT, *Stone Tools as Cultural Markers: Change, Evolution and Complexity*, New Jersey, Humanities Press, 133-146.

GNADEN D., HODAWAY S. 2000, *Understanding observer variation when recording stone artifacts*, «American Antiquity», 65 (4), 739-747.

HOURS F. 1974, *Remarques sur l'utilisation de listes-types pour l'étude du Paléolithique Supérieur et de l'Epipaléolithique du Levant*, «Paleorien», 2, 3-18.

HOVERS E., RAVEH A. 2000, *The use of a multivariate graphic display technique as an exploratory tool in the analysis of inter-assemblage lithic variability: A case study from Qafzeh Cave, Israel*, «Journal of Archaeological Science», 27 (11), 1023-1038.

LAPLACE G. 1966, *Recherches sur l'origine et l'évolution des complexes leptolithiques*, Paris, E. de Boccard.

LEROI-GOURHAN A. 1943, *L'homme et la matière*, Paris, Albin Michel.

LEROI-GOURHAN A. 1964, *Le geste et la parole I - Techniques et langage,* Paris, Albain Michel.

LEROI-GOURHAN A. 1965, *Le geste et la parole II - La mémoire et les rytmes*, Paris, Albin Michel.

LEROI-GOURHAN A., LAMING-EMPERAIRE A. 1966, *La Préhistoire*, Paris, PUF Nouvelle Clio.

LI H.X., YEN V.C. 1995, *Fuzzy Sets and Fuzzy Decision-Making*, Boca Raton, CRC Press.

MARGENAU H. 1950, *The Nature of Physical Reality*, New York, McGraw-Hill.

NICCOLUCCI F., D'ANDREA A., CRESCIOLI M. 2001, *Archaeological applications of fuzzy databases*, in Z. STANÈIÈ, T. VELJANOVSKI (eds.), *Computing Archaeology for Understanding the Past*, Oxford, Archaeopress, 107-116.

READ D.W. 1989, *Intuitive typology and automatic classification - Divergence or full circle*, «Journal of Anthropological Archaeology», 8 (2), 158-188.

READ D.W., RUSSELL G. 1996, *A method for taxonomic typology construction and an example: Utilized flakes*, «American Antiquity», 61 (4), 663-684.

SHOTT M.J. 2000, *The quantification problem in stone-tool assemblages*, «American Antiquity», 65 (4), 725-738.

TIXIER J. 1967, *Procédés d'analyse et questions de terminologie dans l'étude des ensembles industriels du Paléolithique récent et de l'Epipaléolithique en Afrique du Nord-Ouest*, in W.W. BISHOP, CLARK J.D, *Background to Evolution in Africa*, Chicago, The University Press of Chicago, 771-820.

TRIGGER B.C. 1968, *Beyond History: The Methods of Prehistory*, New York, Holt, Reinhart and Winston.

WHITTAKER J.C., CAULKINS D., KAMP K.A. 1998, *Evaluating consistency in typology and classification*, «Journal of Archaeological Method and Theory», 5 (2), 129-164.

WILLEY G.R., PHILLIPS P. 1958, *Method and Theory in American Archaeology*, Chicago, Chicago University Press.

YAGER R.R., FILEV D.P. 1994, *Essentials of Fuzzy Modeling and Control*, New York, John Wiley & Sons, Inc.

ZADEH L.A. 1965, *Fuzzy sets*, «Information and Control», 8, 338-353.

ABSTRACT

It is well known that interpretation always conveys a certain degree of subjectivity, which disappears as soon as interpreted data are stored in a computer database. This may lead to dangerous approximation and possibly to fallacious conclusions. To avoid this oversimplification, it has been suggested to use fuzzy databases, in which attributes may

have a fuzzy nature and be indexed by a numerical coefficient, the fuzzy coefficient, which can be interpreted as the degree of confidence the researcher has in each possible assigned value. This technique has been successfully applied to gender and age assignment for deceased in a cemetery investigation: in this case anthropological data offered statistical parameters that could be used to compute the fuzzy coefficient.

Lithics classification is another field in which fuzzy databases have a potential usefulness, but in this case, no previous statistics may help in determining the fuzzy coefficient. We decided to perform an experiment during a standard typological classification of a flint tool assemblage from Israel. It concerned the classification of 50 tools, by different researchers. Each one was asked to note, besides the typology of each item, an evaluation of the "degree of sureness", or the "possibility" of an item to belong to a particular type, in other words his or her guessed estimate of the fuzzy coefficient.

This paper reports the results of this experiment, in order to evaluate the difference between researchers when performing a classification of tools, to recognize problematic types or items (which mostly differed between the typological lists presented) and eventually to compute a fuzzy coefficient for each type assignment, balancing the different evaluations of experts.