Summary

1. The Interdimensional Sociology as a paradigm for the organization of theory.
2. The syntactic segmentation of sources and the semantic sentence.
3. 'Link-concepts' in the codification of the corpus.
4. The interconceptual links alphabet.
5. The argumentative links alphabet.
6. The segments card.
7. The visual index.
8. The bipolar and tripolar network maps.

This paper does not intend to cover, theoretically, methodologically or technically, the wide variety of specific software used to produce qualitative analysis. Rather, as a sociologist and user of this type of software, my aim is just to suggest some methodological and pedagogical ideas that might possibly improve: (a) the organization of the theory; (b) and the main operational tasks that this kind of software frequently offers to the social scientist in order to organize his research.

1. The Interdimensional Sociology as a paradigm for the organization of theory.

This is an area where a great deal of initiative is granted to the software user. Usually, the core concepts structuring a theoretical frame are represented by a set of 'codes', organized in some coding scheme. The codes often apply to specific segments of texts or other sources to be analyzed, to make possible subsequent searches, different statistics and other kind of data treatment.

These concepts are not previously given in a Classification Language or a Thesaurus by most of these softwares, but they are typically chosen and 'constructed' by the user. Such an attitude may be called open conceptualization, as it is in part based in the 'rhizomatic model' proposed by Gilles Deleuze and Felix Guattari. This unstructured model was applied by Umberto Eco, in the perspective of semiotics, to the systematic conceptual organization of the signification of our social worlds.

Such an open conceptualization may become a good representation strategy of the theory, as each author, in his scientific and life career, often accumulates a

---


2 The forms of this conceptual organization are mainly the 'Porphyrian trees', the 'Dictionaries' and the 'Encyclopedias' (Umberto Eco, 1984, Semiotics and the Philosophy of Language, London, MacMillan Press, pp. 46-86).
rich personal scientific vocabulary, which may symbolize theoretical options or sometimes express conceptual creativity.

Nevertheless, this choice can lead to a lack of communication among social scientists. Note that this conceptual fragmentation problem is not restricted to qualitative analysis, but is associated to the very production of scientific knowledge itself inside social settings, especially in the context of our global information society.

On that account, and as for the organization of a concrete and workable theoretical frame, a practical option has been to use a set of codes considering simultaneously: (a) **legitimated concepts**, that is the ones largely accepted by the global scientific community; (b) and **emerging concepts**, often produced in local academic institutions or inside specific research work, and typically disseminated in periodical publications essays, ‘work in progress’ texts, congress papers, grey literature, etc.

In this mediating perspective (considering both legitimated and emerging concepts), it is advisable for the sociologist to apply, in an early moment, a general and simple conceptual scheme which can cover a great part of reality interpretation, eventually followed by another frame(s) allowing the constant incorporation of new contributions. Thus, here is my first practical suggestion: a **coding strategy can be achieved by taking first a widely recognizable dimension of reality, which can be further connected with other more exigent social dimensions**. Below we can see some of these sociological schemes, to be completed or modified by users inside local or global teams.

In a research, if the sociologist intends to underline the dimension of ‘social levels’ of reality, perhaps this frame can be useful (Table 1):

**Table 1: theoretical codification frame organized by the dimension ‘social level’**

<table>
<thead>
<tr>
<th>1. Macro-sociological level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social processes</td>
</tr>
<tr>
<td>Social structures</td>
</tr>
<tr>
<td>Social classes ...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Mediating social level</th>
</tr>
</thead>
<tbody>
<tr>
<td>(social macro/micro contexts)</td>
</tr>
<tr>
<td>Institutions</td>
</tr>
<tr>
<td>Organisations</td>
</tr>
<tr>
<td>‘Associations’</td>
</tr>
<tr>
<td>Social groups ...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Micro-sociological level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actions / interactions / practises</td>
</tr>
<tr>
<td>Agents / actors / subjects / individuals</td>
</tr>
<tr>
<td>Objects ...</td>
</tr>
</tbody>
</table>

In the case the sociologist prefers to concentrate in the ‘spheres’, ‘realms’, ‘fields’ or ‘areas’, ‘regions’ or ‘continents’ of the social worlds, another (or complementar) option can be this one (Table 2).
These general classes of codes are often used isolated from one another. However, they can be gradually combined or permuted if you consider each of them as a different but interconnected ‘dimension’ of reality. Almost all qualitative software packages allow the simultaneous affectation of several codes to the same source segment. In these conditions, it is easy to associate, to each source segment analysed, one or several codes from a dimension and another code(s) from a second (... n) dimension(s). This strategy establishes a simple but effective connection between the different dimensions, mediated by a segment of a concrete text or other kind of source. Such a process can even lead to a new style of sociological reflection or writing, some sort of Interdimensional Sociology.

Table 2: theoretical codification frame organized by the dimension ’social sphere’

<table>
<thead>
<tr>
<th>1. Socio-economic sphere</th>
<th>2. Political, administrative or power realm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>State</td>
</tr>
<tr>
<td>Leisure</td>
<td>Citizenship</td>
</tr>
<tr>
<td></td>
<td>Science</td>
</tr>
<tr>
<td></td>
<td>Education</td>
</tr>
<tr>
<td></td>
<td>Art</td>
</tr>
</tbody>
</table>

Shortly, when using a qualitative package in theory organization, the choices are unlimited, but it is necessary to begin with some theory ‘shape’ as a reference to code the source segments that constitute our qualitative sampling or corpus.

2. The syntactic segmentation of sources and the semantic sentence

On one hand, the syntactic segmentation of a text or other source means their division in more workable analysis and interpretation ‘formal’ units. In a text, this can be a section or chapter, a paragraph, a sentence, a word or even a letter. The punctuation often helps to determinate the formal units. In my point of view, one artificial and clumsy choice, even if frequently utilized, is to consider a line as an analysis unit. As a matter of fact, this segmentation has little relation with any syntactic structure, even less with any semantic organization, and is only connected to the presentation of a page. The only justification of this choice is the simplicity and ease of management. Nevertheless, in the long term, this can be frustrating for the user, as it makes difficult the interpretation of the text argumentation.

A more useful division may be the paragraph, which is at the same time a clear syntactic or formal unit and a neat semantic or signification unit. Inside the paragraph, a sentence may be delimitated by the punctuation, for example the period, the semicolon, the question mark or the exclamation mark. This assure that a citation can be found by a software search tool while presenting all the words inside a syntactic unit, which allow to interpret more safely its local signification. Nevertheless, it is possible to circumscribe segments through more sophisticated ways, for example
delimitating them through argumentation relationships between sentences or by means of artificial intelligence methods (production rules referring to syntax, etc.)

On the other hand, the semantic segmentation of a source has to do less with the form of the sentences, and more with the units of sense found in these sentences. Such signification units can be identified by complex procedures or methods, for instance an accurate content or discourse analysis or a semiotic reflection. Nevertheless, a growing number of software packages is using more intuitive ways of defining semantic units. For example, often a network model organize theoretical sub-sentences or facts in the following kind of frame: an antecedent concept, a subject term or an independent variable are connected, through a semantic link, with subsequent concepts, object terms or dependent variables. The canonical form here is C1 / R / C2, where C1 is the antecedent concept, R is a relationship or link, and C2 is the subsequent concept. These sub-sentences extracted from formal or syntactic sentences are sometimes called substantive, semantic or logic sentences, as they refer to the underlying and deep signification or the logic beneath their superficial organisation inside English or another national language.

My second suggestion is then to use a syntactic segmentation based mainly (but not exclusively) on the punctuation, articulated to a semantic segmentation that uses both substantive sentences and argumentative links, a notion that we will develop further in this paper.

3. 'Link-concepts' in the codification of the corpus

As mentioned above, the codification – or the indexation, as the information and documentation experts call this operation - connects a code to a source unit, these codes usually representing concepts. On the contrary, a relationship is often used inside a network model, linking typically two concepts. However, and this is my third suggestion, it is possible to represent a relationship by a code, in a similar manner as a concept is associated with a code. In this sense, the relationships should be called link-concepts, as they behave like a concept at least in the codification task. For instance, in Table 3 we have a source on the left and a codification window on the right. We can preview a code to represent the link ‘- -> defines’, and apply this code to the same segment to which we have affected 2 other codes to symbolize an antecedent concept (‘network of relations’) and a consequent concept (‘ethnicity’). The result, inside the codification window, is the assertion ‘network of relations’ -> defines ‘ethnicity’.

Subscribing this strategy has several practical advantages. For example, in this way it would be possible not merely describe, in the codification process, the concept content of our sources but also the relationship content, through a list of codes. Usually, the set of codes instances placed in the codification window underlines only the main theme or concepts. Then, if we make a quick or ‘skimming’ reading, we obtain just a more detailed table of contents or some sort of ‘thematic abstract’.

---

3 For applying network theory to sociological studies, see: Stanley Wasserman; Katherine Faust, 1994, Social Network Analysis: Methods and Applications, Cambridge University Press.
Nevertheless, we could as well construct semantic sentences in the same codification process (that is, sentences constructed by both concepts and links). Like this, we wouldn't limit (or at least concentrate) the use of links to the graphical representation inside a network model, as most software packages do. Furthermore, the network construction is often a manual and slower task, and not an automated operation, like the auto-codification is in some packages. So, we could improve the speed in the definition of links, if this step is done by auto-codification. Moreover, all other software operations that depend on a rich codification (hypothesis testing, statistics, etc.) could become more meaningful if they would be based not only on concept codes but also on link codes.

4. The interconceptual links alphabet.

Let me now precise a little more the question of circumscribing relationships between concepts, or in other words, the construction of interconceptual links. Either inside network graphical representations of the theory, or inside the codification task, it should be useful not to fall in two insufficiencies: (a) first, the deficit of links: sometimes, a software package provides only some basic logic links, like 'is a', 'causes', 'opposes to', etc.; (b) second, the inflation of links: other times, by default, a user can define as many links as the possible number of verbs in a given natural language, like English.

In my perspective (4th proposition), an alternative solution should be a mediating one: it is crucial to develop some kind of interconceptual links alphabet, based in a concise number of logic links, not too restrict and not too longwinded. This alphabet should include, as is already common in the links editor of some packages: (a) the verbal flexions as a synonym of the link name; (b) the relationship sign, which identifies the link in a denotative or objective way, for instance when is useful to do a 'skimming' or superficial and fast reading of the sources; (c) some other link proprieties.
5. The argumentative links alphabet.

This is a feature that complements the precedents. In most packages, a network is built upon the semantic C1 / R / C2 model, where each isolated instance of these 3 elements combination, represent either an epistemological assertion, or a theoretical question, or a typology, or an analytical hypothesis, or an empirical fact.

In these conditions, an argumentation constitute a peculiar reasoning path in the context of the network. Practically, this rhetorical path is normally built by connecting several logical sentences, through a first pair ‘concept-link’, plus a second pair ‘concept link’, and so on. For example, beginning with the substantive sentence C1 / R1 / C2, the concept C2, associated with a new link R2, works as a connector between this first sentence and a second sentence (C2/R2/C3), inside the argumentation process.

Another possibility of a rhetorical path is to connect each logic sentence C1 / R1 / C2 to a second substantive sentence C3 / R2 / C4, through a dedicated argumentative link. For instance, ‘moreover’ is a copulative conjunction used to connect English sentences, which may be represented by an argumentative-copulative link. For the moment, in a practical way, these rhetorical relationships can be defined in a general link editor, inside most software packages. But it should be interesting that a qualitative analysis package could provide an argumentative link editor to the user, with some commands and options of its own, that would activate specific proprieties of the argumentative links (suggestion 5).

6. The segments card.

The next methodological reflection points to the nature of the document and its parts: the sources and its segments could be considered, more deeply, like mediating entities between concepts. What does this mean?

![Fig. 1: a segment card](image)

Usually, in a network map, the concepts act as nodes linked to other concepts by lines, which symbolize the relationships among them. For instance, a star network, often used to organize a sort of concept card, can be described in a metaphorical
way like this: the main concept remains in the centre of the map, behaving like some kind of **sun-concept**, while the others work like satellites or **planet-concepts** around the first one. Clicking on a planet-concept makes it receive the status of a sun term. In this way, it is easy to navigate in a defined conceptual **cosmos**.

Sixth suggestion: however, *we could consider that a document or its segments could perform this central place. For this aim, we may define a sort of segment card, one for each segment in a source, which would complement and communicate with the concept cards*. In Fig. 1, we can see a practical way of doing this, using a star network editor. In this segment card, we can retake the assertion ‘network relations \( \rightarrow \) *defines* ethnicity’, structured by a ‘C1 R1 C2’ logic sentence shape.

However, in this case, the concepts placed by the user in the segment card are not connected directly to each other, but rather joined indirectly, through the node that represents the segment and its source, previously defined as a code. A first connection goes from the antecedent concept node (C1: network of relations) to the segment node, through the new link ‘*defines_inside*’. And a second link connects the segment node to the consequent concept node (C2: ethnicity), by the link ‘*includes_defined*’.

Like this, the links that articulate the concept 1 and 2 to their source are redefined in a way that respects: (a) the semantic relation between the concepts (ex: ‘defines, ‘defined’) and (b) their mediation operated by the segment and its document (ex: ‘inside’, ‘includes’). Thus, we may call them **indirect links or source links**.

Note that this mode of representing segments in a graphical network makes three important distinctions, in a visual way:

1) the difference between the concepts and its sources, these last ones being also presented visually as a node in the network.

2) the distinction between the **legitimated concepts** of the theory, which, in some contexts of the research information system, may ‘dispense’ the presentation of a reference source in a segment card, because they are widely normalized; and the **emerging concepts**, which are not yet recognized as accepted concepts, so their meaning must partially be extracted by the user from a specific source. In other words, even if all concepts are in permanent deconstruction/reconstruction, the emerging concepts can't still be completely understood independently of their source. Thus, the source is present as a node in the network map to help the emerging concept’s interpretation (fig. 2).

3) a last dissimilarity is the one dressed between the codes that act like an antecedent concept and the codes that behave like consequent concepts. This double status of the codes can be perceived directly and visually by their relative situation regarding the segment node (at its left or at its right side, see fig. 1).

7. The visual index.

Seventh proposition: *in the codification process and in the analysis and interpretation phases, it is easier to articulate concepts and their source locations using a sort of visual index, structured by both concept cards and segment cards*. Clicking on the node that represents a segment in a concept card takes the user to the respective segment card, where the segment node takes a central place.
Fig. 2: a concept card

Inversely, while inside this segment card, clicking on the former concept node takes back the user to the concept card where this concept is the main term. Another possible paths are to travel to neighbour concepts nodes (e.g. 'art worlds' in fig.2) or to adjacent segments nodes, while a user is interacting either with a concept card or with a segment card. This way, it becomes practical to navigate visually (that is, using a network view) between concepts and their original documents.

8. The bipolar and tripolar network maps.

Finally, in some software packages, we can find network maps in two main forms: (a) a map with a central concept and (b) a map where a protagonist term doesn't exist. These are precious forms of representing reality and theory. Yet, allow me to offer my seventh suggestion in this regard: in a qualitative analysis software package, a network map editor should be tailored considering the proprieties of the theoretical model that the type of network map will represent.

Let's illustrate this general assumption, as we did above, with a practical example. If we consider Formal Sociology, developed by Georg Simmel, we may interpret reality as a set of social forms. A star network works well in the case of representing a polar institution like the State in a autocratic society, for instance at the national level. In this case, we can easily use a unipolar network (Fig. 3).

---


A sociological expert system using an interesting network of frames to construct everyday situations, inspired in the Goffman's concepts, is ERVING: Edward Brent et alii, 1989,
Fig. 3: an unipolar network (the State in an autocratic society)

But what to do in the case of a dyad (that is, a social form constructed by two actors or by other two social subjects), sometimes concentrating an almost similar power, like it is visible in the competition form, that Simmel has identified inside a sport event or in a war? See fig 4 for an example on competition at the global level, represented by a bipolar network.

And what about the case of a tryad, illustrated by a ‘triangular family’, composed by the father, the mother and their sons? Fig 5 gives an example of a family at an institutional and micro-social level, in a tripolar network. As you see, each pole representing one member of the family has its own tree of social branches, and can ‘take the lead’, instead or simultaneously with the other protagonists, in a sort of triangular intra-network working inside a more complex set of relationships, the surrounding family meta-network.

All this means that these different social geometric forms could be better symbolized not only in unipolar networks, like a simple star shaped network does, but as well inside bipolar or tripolar networks.

Fig. 4: a bipolar network (two of the main actual society paradigms)

Fig. 5: a tripolar network (the triangular family in a social proximity context)

Let me conclude with this comment: to study our contemporary societies, we can use infinite methods and strategies inside scientific work, to improve the 'democratic' nature of theory, as a non-unipolar network may suggest, in its 2 mains shapes: the non-polar network or the multipolar network.