Contributions to Methodology and Statistics
A. Ferligoj and A. Kramberger (Editors)
Metodološki zvezki, 10, Ljubljana: FDV 1995

Sets and Networks: Two Research Approaches[1]

Jan Makarovič[2]

Abstract

Any phenomenon may be described either as *set* or as *network*. Set refers to description of *elements* of a given whole, while network refers to their *connections*. If every element in a set is equal to any other, we have to do with a *full* set. If no element in a set is equal to any other, we have to do with an *empty* set. Likewise, if every element in a network is connected with all others we have to do with a *full* network; if none of them is connected with any other, we have to do with an *empty* network.

Durkheim's theory of social evolution is a classical instance of investigation into the correlations between set and network aspect of reality, although Durkheim himself was not aware of it. According to him, social solidarity (a special kind of social network) may be dependent either on *similarity* (full set) or upon *differences* (empty set) between people. In the first case, we have to do with *mechanical* solidarity, in the second with *organic* one. Similar considerations may be met in Spencer, Marx, and others. The explicit use of the logic of sets and the logic of networks may clarify them and prevent logical misunderstandings that are rather frequent in sociological theory.

1 Hume about associations and causality

As far as I know, the difference between sets and networks, which I am going to cover in this paper, was nowhere better demonstrated than in Hume's **Treatise of Human Nature,** written more than 200 years ago (1739-1741). David Hume, the greatest British philosopher, is known first and foremost as a **sceptic**. He is one of those great minds who became famous not because they have shown what they

[1] In honour to the 100th anniversary of Durkheim's work "De la division du travail social" (1893).

[2] Faculty of Social Sciences, University of Ljubljana, P.O. Box 47, 61109 Ljubljana, Slovenia

know, but because they examined human experience critically and demonstrated what man **doesn't** know, while he thinks he does. He distinguished clearly between what really **is** in our experience, and what is only an illegitimate inference from it.

What, then, is the **actual** content of our experience, according to Hume?

He claims our experience consists of **perceptions**. Perceptions are divided in two groups: the primary and the secondary. The primary group of perceptions consists of **impressions**, the secondary of **ideas** -- nothing but fainter reflections of impressions. Naturally, ideas are frequently obscure and ambiguous: one must have recourse to associated **impressions**, in order to clarify them. If no such impression can be found, it is obvious that we are not dealing with an idea at all, but with an **empty word**. In other words, we **do not know** what we are talking about, and in that case the "idea" in question must be discarded.

Ideas are not distributed randomly, but form greater or lesser, more or less compact groups. The structuring of ideas is based on **associations**. We have seen that ideas are based upon **impressions**. Likewise, associations between ideas are based upon **relationships between impressions**. According to Hume, there are only two kinds of such relationships: **resemblances** and **contiguities** in time and space. On the basis of these, two kinds of associations are formed: (a) such as between a man and the painting on which he is portrayed; (b) such as between St. Denis and Paris (Hume's own presentation of his **Treatise**).

At first sight, these epistemological reflections seem a little naïf. Everybody knows that human **thinking** is something different from a flow of dreamer's **associations**. However, it must be conceded that **classification** is one of the most important foundations of human thinking, and that it is based upon establishing of **resemblances** and **contiguities**. If I say: "Socrates is mortal", I include the entity "Socrates" into the set of mortal beings, and this is done upon the basis of **resemblance** between Socrates and other mortal beings in respect of mortality. If I say: "The Bled Castle overlooks the Bled Lake", this expresses a spatial contiguity between the Bled Castle and the Bled Lake. By such statements, I can completely **describe** the world I live in. However, I cannot **explain** it.

Here Hume's main epistemological problem appears. He is aware that resemblances and contiguities of impressions are given empirically and that on the basis of them it is possible to present an integrated picture of our world. Nevertheless, it is impossible to **explain** the phenomena on this basis, because it is impossible to infer **causal** relationships from resemblances and contiguities.

To be sure, Hume presents a **psychological** theory of how our notion of causality is formed. However, this theory, being psychological rather than not logical, can only explain our **belief** in causality -- not causality itself. This criticism of causality is the basis of Hume's **scepticism**. It awakened **Kant** from his "dogmatic slumber", as Kant himself admitted, and has given the first impetus to his famous **Critique of the Pure Reason**. Kant started from the statement that

it is impossible to arrive at the notion of causality empirically and that therefore this notion must exist as a "pure form" in the mind itself. This is the "Copernican turn", by which Kant ended the era of English classical philosophy that had been based on empiricism, and introduced the era of German classical philosophy that discovered the **creativity** of the mind.

We will not dwell upon these philosophical debates. We only want to stress the crucial epistemological difference between the principle of **association** that is the basis of the **description** of the world, and the principle of **causality**, that is the basis of its **explanation**.

2 The logic of sets and the logic of networks

Aristotle's logic is basically a logic of **classification**. Take, for instance, the statement we have mentioned before: "Socrates is mortal". In this statement, "Socrates" is the **subject**, and "mortal" is the **predicate**. Thus we get the formula "S = P": subject S has the predicate P. Basically, this is the **inclusion** of S into the class P. The fact that we are dealing with a logic of classification may be graphically demonstrated by Euler's circles.

Cantor's **set theory** can express such relationships much more precisely, by the formula $S \cap P$. By that move, the ambiguous term "is" is avoided ("Socrates **is** mortal"), the Aristotelian **copula** producing the illusion of **identity** that has been since ever the main source of scholastic puerilities, from Saint Anselm of Canterbury till Hegel. On the other hand, set theory understands the relationship of inclusion as only one of the possible relationships and is therefore much superior to Aristotelian logic that cannot express adequately many statements, such as, for example, "Bled is smaller than Ljubljana," or, "Bled lies in a northwestern direction from Ljubljana."

Set theory is superior to Aristotelian logic as well as to elementary arithmetic, that uses only the relationship of **equality** (=). If I say "Socrates is mortal", I do nothing but establishing the **equality** between Socrates and other mortal beings as regards the mortality. Thus in a way I associate Socrates with other mortal beings. In a similar way, in computing mathematical equations, I associate the left side of the equation to the right, as being equal to it.

Set theory, on the contrary, uses other relationships as well. However, these relationships are ultimately nothing but **comparisons**. If I say Bled is smaller than Ljubljana, or that it lies in a northwestern direction from it, I thereby only **compare** Bled to Ljubljana, and **describe** it in its relationship to the latter. I do not **explain** Bled, as I would if I discussed the formation of its glacier valley, its early settlement, the foundation of its castle, and so on. My statement does not imply any **"real connection"** between Bled and Ljubljana, to use Hume's terminology. Basically, comparisons remain on the same level as equations, and

equations are but a special category of them. On the other hand, all comparisons of elements within a given set can be expressed as equations, although they are thereby simplified. The reason is that all comparisons, except the comparison of equality itself, imply **non-equality**. If I say that Bled is smaller than Ljubljana, this implies that it is **non-equal** to it. Thus, all such statements may be divided into two groups: in the group that says what Bled is like, and the group that says what Bled is **not**. Now we are in the Aristotelian logic again.

An entirely different approach is implied in **network theory**. This theory is not primarily interested in properties of elements and in comparisons among them, but in **processes** that take place among them. These processes cannot be simply identified with **causal** processes. In another place (Makarovič, 1993) we have distinguished four kinds of such processes. Causal processes are only one kind. The essential thing is that something real **happens** among the elements, so that we are not merely comparing them. Needless to say, such processes correlate with specific properties of the elements. Nevertheless, as Hume has warned us, they cannot be simply deduced from such properties. Here we have to deal with two different aspects of reality, and with two different research approaches.

This does not mean that they should not be combined. On the contrary, any satisfactory research work must combine them. However, before such a combination is possible, a clear formulation of specific problems arising from both approaches is necessary.

In the next chapter, we will cover one such problem: **formal classification** of sets and networks.

3 The formal classification of sets and networks

In any research endeavour, classification is the first step. We must clarify the categories we are dealing with, and distribute our phenomena among these categories. Naturally, we must first ascertain the most general categories, and only then proceed to the more specific ones.

I think that sets and networks are the most general categories in any research work. If this is true, then we must begin with the classification of sets and networks. Needless to say, here we do not have a classification of elements of sets and networks (by the way, network theory is not interested in such classifications at all), but only a classification of sets and networks as **wholes**.

Naturally, we renounce any ambition of classifying the content of sets and networks. Taking into account the infinite variety of elements of sets and networks that exist in the world, and the infinite variety of their relationships, would be presumptuous. However, if we succeed in defining the relationships that appear in sets and networks on the most general level, it may be possible to classify the whole entities simply on the basis of the presence or absence of such relationships.

Of course, such a classification will be entirely **formal**, as it will be abstracted from any empirical content.

Our first problem is the distinction between relationships that appear in sets and those that appear in networks. In sets we deal only with **comparisons** between elements, while in networks we must reckon with **processes** that take place between them. However, the English language does not have terms that would cover each of the two meanings. We must resort to the German language. In German, there are two suitable terms, although they are frequently used rather carelessly: **Beziehung**, implying a process taking place between two elements, and **Verhaeltnis**, referring to comparisons between them. In Slovene, corresponding terms would be **odnos** and **razmerje**. We propose as English equivalents the terms **connection** and **ratio**. To be sure, "ratio" usually implies only a quantitative relationship, but we will use it in the qualitative sense as well. Consequently, it may refer not only to values of cardinal variables (f.i., 1:2), but to values of nominal variables as well (f.i., yellow:orange).

There may exist **quantitative** differences between connections and ratios. If two soldiers fight desperately to the end, the connection between them is strong; if they try to avoid the combat, their connection is weak. Likewise, if two lovers are madly in love, the connection between them is strong; if they are less emotional it is weak. The ratio 1:2 is higher than the ratio 1:4, because in the first case similarity between both numbers is greater. Likewise, man is more similar to another man than to a woman, the orange is more similar to red than to blue, and so on.

Such quantitative differences may be simplified, if we presuppose only two values for each connection and ratio (either **exist** or **not exist**). Thus, a **similarity** between two elements (by "similarity", we understand both "resemblance" and "contiguity", in Hume's terms) or be absent. On the other hand, a **connection** between two elements may exist or not exist.

In spite of this simplification, an immense variety of relationships may still appear within a given set or network. Within a set of only four elements, A, B, C, and D, for instance, A may be similar to B and C, but not to D, or it may be similar to C and D, and so on. Therefore, we must simplify again, and take into account only the **quantity** of relationships, not their specific configurations. Furthermore, we admit only three quantitative levels: the **general**, the **particular** and the **individual**. Relationships appear on the general level when **all** elements are interrelated. They appear on the particular level when **some** of them are related to some others. Finally, they appear on the individual level when **none** of them is related to any other.

This general rule may be applied to sets and to networks alike. However, the result will be quite different for sets and for networks (Figure 1).

1. **For sets**: A greater or lesser extent of **similarity** may exist among the elements of a set. When all the elements of a set share the same characteristic, the set may be labeled **homogeneous**; when some of them do, it may be labeled **heterogeneous**; finally, when every element is different from every other, it may be labeled **unique**. We speak about "uniqueness" of sets although **stricto sensu** not sets but their elements are unique in this case. However, we couldn't find a more appropriate term.
2. **For networks**: In a similar way, networks may be distinguished on the basis of the extent of their **connections**. Networks, in which everything is connected with everything else may be labeled **integrated**; if some elements are connected with some other elements, they may be labeled **segmented**; finally, if nothing in them is connected with anything, they may be labeled **fragmented**.

Thus, three logical levels - the general, the particular, and the individual - are reflected in specific ways in sets and in networks

LEVEL OF INCLUSIVENESS	SETS	NETWORKS
• **General**	homogeneous	integrated
• **Particular**	heterogeneous	segmented
• **Individual**	unique	fragmented

Figure 1: A simplified classification of sets and networks

The translation of logical levels into the language of sets has evidently a quite different meaning than their translation into the language of networks. Thus, "individualisation" in society may mean:

- **either** that everyody is **isolated** from others and is not dependent upon foreign influences (network logic);
- **or** that everybody is **different** from others (set logic).

All sets and networks fluctuate between two ideal states: the state of extreme **richness** of relationships, and the state of **no relationships** (individualised state). The state of extreme richness of relationships may be labeled "full" state. In this sense, homogeneous sets may be labeled **full sets**, and integrated networks may be labeled **full networks**. On the other hand, unique sets may be labeled **empty sets**, and fragmented networks may be labeled **empty networks**.

Thus, sets and networks may be understood as **variables** with value 1 on the full side and value 0 on the empty one. On that basis, **correlations** between sets and variables can be computed.

What these correlations be -- positive, negative, or zero? According to Hume, it is illegitimate to infer causal relationships from resemblances and contiguities as such. Therefore, we should be cautious in expecting a **general** correlation between the extent of similarities on the one hand and extent of connections on the other. The question of such correlations is an **empirical** question, and the answer to it depends upon the **specific nature** of the phenomena we are observing.

Therefore we will limit ourselves to a specific research field of **general sociology**. I will state some reasons to expect that in this case the correlation in question will be **negative**: the more developed the set, the less developed the network, and vice versa.

The question of the origin of human society is the basic sociological question. More than two thousand years ago, this question was clearly formulated by Plato in his **Republic**. Plato's answer is that human society is based on **differences**. People do not seek the company of other people who are similar to themselves and can perform the same tasks as themselves can, but of those who are different from themselves and can therefore compensate their own shortcomings. Therefore, a man seeks a woman, a weak man seeks a strong man who can help him, a dull man seeks a wise man who can counsel him, a peasant seeks a smith who can make his plough, while the smith seeks the peasant from whom he may obtain food in return to his plough, and so on.

Plato takes the differences among people for granted and understands society as a consequence of these differences. However it must be stressed that society is not only the **result**, but also the **cause** of differences among people. Precisely because people are united in a society, they develop the need for division of work, and consequently for specific social rôles. When Robinson lived on his island alone, he had to perform all social functions himself. Suppose that at the same time another Robinson lived on another island: he too would need perform the same functions. Precisely because our two Robinsons were **separate**, they had to be **alike**. On the other hand, when Friday came to Robinson's island, he became his **slave**, and Robinson became his **master**. Their functions became different precisely because they were united. Thus, in society integrated network is associated with unique set.

However, this problem is more complicated than this. The man who did most to elucidate it was Emile Durkheim.

4 A few comments on Durkheim's theory of social solidarity

In 1893 Durkheim published his first great work, **De la Division du Travail Social**. This was one of the most important sociological books ever published,

because it stated -- more clearly than most others -- the basic sociological question: **how is human society possible?**

Durkheim gives **two** answers to this basic question. One of them is basically identical with Plato's answer: society is based on **differences** among people, and upon their **division of labour**. This is the basis of their social **solidarity**. However, Durkheim reminds us that such a basis of solidarity is possible only in **modern** societies with a highly developed division of labour. The question arises, then, how is society possible in **lower** developmental stages, when the division of labour is less developed?

In these developmental stages, according to Durkheim, solidarity is based on the **opposite** principle - **similarity**. People feel solidarity with others simply because they feel themselves similar to them, belonging so to say to family.

However, is that true? Some twenty years after the publication of Durkheim's book, the great anthropologist Bronislav Malinowski did something no European had done before: he lived two years uninterruptedly among the Trobriand islanders in the Pacific. There he once had an interesting experience. He casually remarked, in a conversation with a native, that he looked very similar to his brother. The man became as pale as a native from the Southern Seas can, and went away. Only later Malinowski learned that speaking about one's similarity to one's relatives is the worst **insult** possible on the Trobriand islands. The colloquial terms connected with this are the same as those designating one's sexual intercourse with one's own sister, which is also the most abject conduct possible (Malinowski, 1926). Here Freud's theory about the origins of the prohibition of incest comes to mind. Erotic bonds among members of the same family would cause **jealousy** among them, and disrupt the solidarity in the family. In this sense, being **similar** to one's brother means that you want to possess **the same sister** as he. Such dangerous feelings should be avoided and are therefore considered as abject (Girard, 1972: 92 ff.).

In his **last** great work, **Les formes élémentaires de la vie religieuse** (1912), Durkheim clarified his view about solidarity in primitive societies, such as that of the Australian aborigines is. It became clear that solidarity in such societies is not based upon their psychical or physical similarity as such, but upon the common identification of their members with a **totem**, or any other **common symbol**. A society without developed forms of cooperation and division of labour **must** base solidarity among its members on **symbolic** representations. Not on external similarities among its members, but on the **distinction** between the material and the symbolic levels. The members of a primitive society **are** similar just because they are identified with a common symbol. Likewise, their solidarity is not **immediate**, does not spring from their similarities or differences as such, but is **mediated** through the symbol. Members of society are similar, form a homogeneous set, not because of their immediate similarity but because of their common participation in their symbolic world. They are integrated indirectly, not

directly. This is diametrically opposite to the modern society where solidarity is based on direct, not indirect connections.

Thus, Durkheim distinguishes between two kinds of social solidarity. The first he designates as **mechanical**, because here individuals are attached to society in a mechanical way, in the same manner as things are attached to persons: "The individual conscience, considered in this light, is a simple dependent upon the collective type and follows all of its movements, as the possessed object follows those of its owner" (Durkheim, 1933: 130). We have seen that "society" here appears in the form of **symbolic representations** to which individuals are directly attached. On the other hand, they communicate among themselves only **indirectly**, via these symbolic representations.

On the other hand, **organic** solidarity is based on differences between people, upon their specific functions and their complementarity, like the functioning of a living organism is based on the complementarity of its organs. This theory of two forms of solidarity may be understood in terms of sets and networks. Solidarity is basically a form of social integration, and refers therefore to the network aspect of society. This network aspect corresponds to the set aspect of society: in primitive society integration is based on homogeneity, while in modern one it is based on heterogeneity. In other words, correlation between the set and the network aspect of society is positive in primitive society, and negative in the modern one.

Generalizing still further, it is possible to conclude that correlation between set and network aspect the society is positive where integration is based on indirect connections, and negative where it is based on direct ones (see Figure 2).

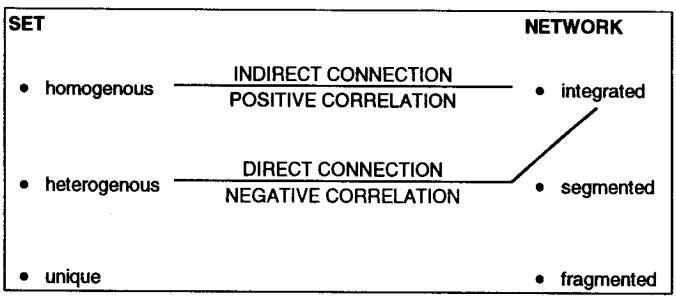

Figure 2: Correlations between the set and the network aspect of society, inferred from Durkheim's theory of social solidarity

However, Durkheim discussed the problem of social integration solely as a problem of **solidarity**, but it must be stressed that solidarity is the basic mechanism of social integration only in **classless** societies. In **class** societies, integration is based on class **antagonisms**. Like Durkheim, who presents two kinds of social solidarity, Marx distinguishes two kinds of class antagonisms

(Capital I. 1.1.4). In the preindustrial class society, social integration is based on one-sided **personal dependencies** that are of course based on class **differences**. Industrial society, on the other hand, is based on the **equal exchange** of commodities. Everyone is obligated to give an **equal** value of a commodity, in return for the value received. In regard to commodities, people are equal; in the same way as in mechanical solidarity they are equally attached to a common symbol. But their equality is not immediate; it is mediated through the commodities. Society itself is **fragmented**, as is very well expressed in the title of Riesman's famous book, **The Lonely Crowd**. Everybody is equal in the market (as well as before the law and before God), but this equality obtains among individuals that are as such unequal. Their inequality as individuals is just another side of their equality in the market. Industrial society is homogeneous, in the mechanical, not in the organic way. This is similar to mechanical solidarity, where individuals are different as such, but equal as regards their common identification with a symbol. However, it must be borne in mind that in the latter case we are speaking about qualitative similarities and differences, while in industrial society (antagonistic and class society), (in)equality is essentially quantitative. In the same way, quantitative (=class) differences are essential in the pre-industrial class society, based on one-sided personal dependencies.

On the other hand, industrial society is fragmented, according to Marx, because unequal individuals can appear as equal partners in the market only if they are independent individuals.

Thus, industrial society as understood by Marx is strangely similar to Durkheim's mechanical solidarity. However, it is not solidary, but antagonistic. Therefore we could label it mechanical antagonism.

Interestingly, the same basic social regularity may be inferred from Marx' theory as from Durkheim's: where integration is based on indirect connections, it is associated with homogeneity; where it is based on direct ones, it is associated to heterogeneity. indirect connections cause negative correlations between set and network aspect of society, while direct connections cause positive relations between them (see Figure 2). This general regularity is the more striking because it follows not only from theories of two radically different authors, but refers to quite different types of society.

On the other hand, communist society, that should replace capitalist society according to Marx, is similar to Durkheim's organic solidarity. While Marx' model of communist society is utopian and never very precisely formulated, Durkheim's model of society of organic solidarity is sociologically sound and carefully drawn. It could be understood as an operationalization of Marx' communist society "in which a free development of everybody is the condition of free development of anybody", as he put it in his Communist Manifesto.

From a contemporary (postindustrial) perspective, information society may be understood as a transition from mechanical antagonism, as understood by Marx, to

organic solidarity, as understood by Durkheim. This transition is represented schematically in Figure 3. There is no place for elaboration of processes involved in this transition in the present paper; however, it cannot be doubted that precisely this is one of the crucial theoretical, methodological, and empirical challenges of contemporary sociology. We have seen how can our distinction of sets and networks be helpful for this task.

FROM:	TO:
• mechanical	• organic
• indirect connections	• direct connections
• antagonism	• solidarity
• competitiveness	• cooperation
• homogeneity	• uniqueness
• "cocacolization"	• individualization
• positive correlation between sets and networks	• negative correlation between sets and networks
• stress on quantity	• stress on quality
• quantitative (in)equality	• qualitative differentiation
• stress on material production	• stress on information processing
• exchange	• communication
• materialist values	• spiritual values
• etc.	• etc.

Figure 3: Some aspects of the transition from the industrial to the postindustrial society

Formalization and operationalization of this transition, as it appears in various spheres of social life, is one of the great methodological challenges of contemporary sociology.

5 Conclusion

In this paper we stressed the utility of the distinction between sets and networks for sociological theory and methodology. We applied this distinction to the basic sociological question - the question of social integration - and thus encountered one of the greatest minds in sociology, Emile Durkheim. Although we think that his theory is insufficient, it was a real break-through in sociology. We think that this theory gets a new vigour if looked in the light of the distinction between sets and networks.

Our reflections are of more than theoretical interest. Mankind is now living in the era of transition from the industrial to the post-industrial society. Translated into the language of our theory, this would mean that we are undergoing a transition from **mechanical antagonism** into **organic solidarity**. We have to deal with quite different kinds of sets and networks in the two kinds of society. The transition between these societies is a highly relevant question for understanding contemporary society, and the planning future social development.

Naturally, abstract and greatly simplified models, such as are used here, cannot do justice to concrete social situations. However, they may be useful starting points for research.

Unfortunately, space allows only a rough sketch of our theory. However, we have discussed the same issues from various perspectives elsewhere (see, for instance, Makarovič 1993 and Makarovič 1995) and are referring the interested reader to them.

References

[1] Durkheim E. (1933): *The Division of Labor in Society*. New York: Free Press.

[2] Girard R. (1972): *La violence et le sacré*. Paris: Grasset.

[3] Makarovič J. (1993b): *Logika dela: zgodovina in prihodnost* (The Logic of Work: History and Future; English summary). Ljubljana: Znanstvena knjižnica Fakultete za družbene vede.

[4] Makarovič J. (1994): *Grenki kruh* (The Bitter Bread; English summary). Maribor: Obzorja.

[5] Malinowski B. (1926): *The Father in Primitive Psychology*. London.