
MAN AND SOCIETY: NESTED INTENTIONALITY

In the introductory chapter I described my theory of categories as, among other things, an attempt to transcend the opposition between atomism and holism. I shall now repeat and summarize what is of immediate interest in relation to such a distinction before tackling the main theme of the present chapter: nested intentionality.

15.1 FOUR CONCEPTS OF 'PARTS'

As regards the question of atomism and holism, one can begin by distinguishing four different basic senses of 'part'. The first three are: (1) part as *element* in a set, (2) part as *moment* of a whole, and (3) part as a *spatio-temporal part*. Part-whole relations in the first sense lack ontological interest and are mentioned here only for the sake of completeness. The element-set relation is *either* a purely linguistic creation as when one constructs a set of completely arbitrary objects, for example the set consisting of the sun, the first word on this page, and the chair I am sitting on. Nothing other than their being members of the same set unites these three elements in a 'whole'. *Or* the relation may be based on something which truly exists and which unites the elements in question, for example a certain property such as when one constructs the set of all red things. Ontologically interesting whole-part relations can exist here, but they concern the universals used to construct the set, not the actual relation of belonging to a set.

That something is part of a whole in the sense of a moment means that the part cannot exist without being a constituent of a more comprehensive whole. Such a part (aspect) is thus not, as atomism would have us believe, existentially primary in relation to

the whole. On the other hand, neither can the comprehensive whole in question exist without (at least some of) its moments. Part and whole depend on one another. In this sense even Democritus's indivisible atoms have parts. A spatially indivisible corpuscle/atom has as moments (parts) both shape and volume. Without these parts, the atom could not exist. There is no absolute atomism. No matter how the atoms in an atomistic world view are specified, they must contain moments, i.e. *one* sort of part. It is true that it is only philosophers within some traditions who have been happy to talk of 'moments' as 'parts' but this is a merely verbal matter. The important point is that atoms and moments are connected together.

There are many sorts of parts in the third sense, spatio-temporal parts. One thinks first, perhaps, of aggregates of things. One can also distinguish spatio-temporal parts amongst instances of single universals. Most interesting in this context are the inclusive universals example. We must also remember the distinction between parts and lower-dimensional parts. Normal parts are parts with the same number of spatio-temporal dimensions as the universal in which they are included; lower-dimensional parts are those which have fewer dimensions. With regard to normal parts the whole is one-sidedly existentially dependent on the included parts. The situation is more complicated as regards lower-dimensional parts. In at least certain cases there is a relation of mutual existential dependence: two-dimensional surfaces cannot exist other than as lower-dimensional parts of three-dimensional volumes, and of course three-dimensional volumes cannot exist if they do not have two-dimensional surfaces as lower-dimensional parts; they have to be limited by such surfaces (see also the discussion on page 88f.). Patterns and Gestalten have spatio-temporal parts and addition can be defined even for these parts. There is no reason to employ the expression 'the whole is more than the sum of its parts' in connection with patterns and Gestalten in so far as the parts in question are equi-dimensional parts. When one looks at lower-dimensional parts then it is always true that 'the whole is more than the sum of its parts'; lower-dimensional parts cannot, by definition, via *addition* result in a whole with an increased number of dimensions.

The claims I have put forward about parts in the sense of moments, and parts in the sense of normal and lower-dimensional

spatio-temporal parts, go beyond what traditional atomists usually accept without, for this reason, coinciding with what holists usually say. Atomists usually do not accept, for example, internal relations, i.e. mutual existential dependence, and holists usually do not accept one-sided existential dependence. And neither atomists nor holists have clearly understood the existence of inclusive universals. In spite of this, what I have said so far does not lead to a *radical* break with the atomistic viewpoint. Spatio-temporal parts are ontologically primary in relation to that type of spatio-temporal whole (things, aggregates of things, machines, organisms, and natural processes in general) which they constitute. The radical break comes with the existence of parts in the fourth – as yet unnamed – sense.

Parts in the sense of moments are based on the category of existential dependence, and spatio-temporal parts are based on the category of container space. Parts in the fourth sense, however, are based on the category of intentionality. The whole of which such parts are parts is an intentional whole with the two features: 'pointing' and 'connection at a distance'. Parts in the fourth sense I shall call 'participants' – these parts participate in a whole.

One example of parts of this last-mentioned type are the members in an organization. To be a member of an organization is not the same as (1) being an element in a set, or (2) being a moment, or (3) being a spatio-temporal part of an organization. One can of course construct the set of all members of an organization, but this set is not identical with the organization. An organization exists in space and time, while a set is an abstract entity. To be a member of an organization cannot, either, be the same as being a moment of the organization, since the organization can persist even if one member has died, and vice versa. Both the members and their organization exist, in space and time, but this does not imply that the relation 'to be a member of' is identical with the relation 'to be a spatio-temporal part of'. The spatial inclusion relation is, for example, transitive, but the membership relation is not. If a is a spatial part of b , and b a spatial part of c , then a must be a spatial part of c . But if a is a member of (the organization) b , and b is a member of the organization c , then it does *not* follow that a is a member of c .¹

One thus sees, as a result of simple considerations, that the membership relation is a very specific part-whole relation.

Similarly, it is easy to see that the intentionality category here plays a decisive role. In order for someone to be a member of an organization, he must be seen and conceived as a member, if not by himself then by *someone*. Membership without some sort of intentionality does not work. The intentionality category can, under certain conditions, and like the space category and the existential dependence relations, function as a 'unifying category'. Just as existential dependence, by linking substances and properties, gives rise to the new category of state of affairs, so too, as we shall soon see, intentionality gives rise to nested intentionality.

There is intentionality of two kinds: presentational and representational. It is presentational intentionality which forms the basis of the most interesting part-whole relations, but representational intentionality also gives rise to special types of wholes. Assume that five people who know each other but live on different continents think of each other at the same time. Person *A* thinks of *B*, *C*, *D*, and *E*; person *B* thinks of *A*, *C*, *D*, and *E*, and so on. Obviously some sort of unity among the five persons is created; a unity that goes beyond the external spatial relations which naturally connect them and make them constitute a pure spatial pattern. The example is trivial in the sense that normally we encounter similar situations all the time. It obtains philosophical significance through the analysis of the intentionality category that I have carried out. I would like to stretch Leibniz's terminology a little and say that representational intentionality can give rise to 'monadological wholes'. In the above example we have five persons (with spatial bodies, that is to say, they exist in container space), all of whom have an intentionality which is founded upon their bodies. Each intentional state or act points to the other persons, but each intentional phenomenon is spatially distinct from the others. Each person becomes a monad that mirrors the others.

The situation is quite different in the case of presentational intentionality, since I have argued that in such cases we can be in direct contact with the intentional goal. Intentionality and its correlate here make up a unit and a wholeness in space that goes far beyond purely spatial relations; we get a much stronger type of connection than that which gives rise to 'monadological wholes'.

15.2 AGENTS AND SUBJECTS

The whole of which participants are parts is always a whole consisting of at least two participants who are subjects. A subject is a spatio-temporal whole consisting of a body which has, or has a disposition to have, intentional acts (cf. section 13.1). Before commencing on a discussion of this concept of 'participant', a few points should be made regarding both the subject category as such and the intentional connection as such. We shall analyse some intentional phenomena which only include *one subject* and an intentional correlate which is an *object* or a naturalistic state of affairs.

Certain emotional and intentional states, like happiness, sadness, and fear can have naturalistic states of affairs as intentional correlates, and the intentionality in question can be presentational. I can be happy about the sunrise, sad about the rain, and frightened about the avalanche which is approaching me. The sunrise, the rain, and the avalanche exist independently of my perception of them, but they nevertheless become connected to me – and I to them – through perception. We become connected by intentionality.²

The claim that these states are intentional does not contradict our earlier analysis of states in general (sections 7.3 and 11.2). A state of happiness, sadness, or fear contains as do other sensory states, a tendency to carry out certain actions. Emotional states like intentions (see pages 106–7 and 208–9), contain *both* intentional acts *and* tendencies to act in certain ways.

Subjects who are in emotional states like happiness, sadness, and fear exist in the space-time taken up by their substrata. That which the emotional state is *about* is, in the case of presentational intentionality, in some sense a part of the state; but due to the intentionality 'hop' in space, it would be mistaken to call it a spatio-temporal part of the subject. The most adequate description possible is to say that the subject and the naturalistic state of affairs (= the intentional correlate) are connected by intentionality.

Normally we understand emotional states as something we are affected by, not as something we choose. One is not an *agent* simply because one has conscious experiences of happiness, sadness, fear, and such. The agent category does not coincide with the subject category; agents are a special sort of subject. These subjects,

however, play an important role in what follows and so a few words should here be said about the agent category. It is best understood by means of the concept of rationality and rational action. Like happiness, sadness, and fear, rationality involves intentionality. Whether it also includes tendencies and is a sort of intentional *state* is a question which will have to be set aside for the moment. To be rational is to be rational *in* a situation or to be capable of being rational in a situation. The one who is rational must have his situation as an intentional correlate. In order to make the issue more precise, I shall return to chapter 5 and the presentation of methodological individualism. That I make use of this ‘-ism’ as an example is of no great consequence. The type of explanation pattern which I presented under the label ‘methodological individualism’ can be found in such widely disparate traditions as neo-Kantianism, hermeneutics, existentialism, and marxist praxis-philosophy. Here, there is a hidden meeting-place for apparently contradictory traditions. But this is not strange. All these traditions take as their starting point the everyday category of agent, i.e. the fact that we often act after having made a choice in a situation.

Noretta Koertge has suggested, as a first approximation, a model of methodological-individualistic explanations.³ I shall not reproduce the model in order to comment on Koertge’s analysis, but because the model provides a good background to the remarks I should like to make. This is how the model looks:

- (1) *Description of the situation*: Agent *A* was in a situation of type *C*.
- (2) *Analysis of the situation*: In a situation of type *C*, the appropriate thing to do is *d*.
- (3) *Rationality principle*: Agents always act appropriately to their situations.
- (4) *Explanandum*: (Therefore) *A* did *d*.

In the model the *explanandum* is a logical consequence of the first three points. I shall go through them one at a time, and begin at the bottom with point three. In the formulation that the principle of rationality has received here, it is obviously false. Normally we are affected by diverse factors which often lead to our irrationally diverging from the reasonable action we had decided on. Agents only seldom act completely adequately in relation to their

situation. Many of the philosophers and scientists who have advocated this model of explanation have pointed out how frequently such irrational divergences occur.⁴ Rational actions, they have pointed out, must be seen as an ideal type. The above model must be supplemented with a clause saying that no irrational factors play a role in the situation in question; otherwise the *explanandum* cannot be derived.

The principle of rationality also contains an unclarity. As it has been formulated, it leaves a categorial question open: Do agents always act rationally because they have a *tendency* to act that way (a tendency which can then complement or counteract irrational factors); or do they always act rationally because they *spontaneously* (in the sense of the spontaneity category of chapter 7) *choose* rational actions? In the formulation given here, the rationality principle says only *that* agents act rationally; but their actions must be either spontaneous or occur *causa sui* and be conditioned by an inertial tendency. In the latter case rationality can be put on equal footing with whatever intentional state one pleases. Rationality becomes a special kind of emotional state. The majority of those who have advocated a rationality principle seem to have had the former case in mind. One takes for granted that the agent *could have acted otherwise*. He could have acted otherwise by choosing to try to realize another intention, and that choice or decision is to be understood as a change which falls under the spontaneity category. I shall quite simply define the agent concept in such a way that an agent is a subject whose intentions (tryings) could spontaneously have been different. Here, everything turns on the category of spontaneity that was introduced in section 7.1. An agent is such that he can have a spontaneous intention to do something. This means that it is impossible to predict with certainty an agent's intention: one may only speak about normal intentions. Of course, these intentions also have a tendential side, which explains the realization of the actions and makes possible their combination with irrational factors.

The rationality principle should therefore be reformulated in the following way:

(3) *Rationality Principle:*

- (a) Agents normally decide to act in a way appropriate to the situations in which they find themselves.
- (b) Agents can be affected by irrational factors.

As point (2) (*Analysis of the situation*) has been formulated, it presupposes that the rationality of the situation is independent of the agent. No agent is even mentioned. But this must be wrong. Let us once again look at the classic example of the consumer in the market – more particularly, in the market described earlier (see page 62ff.) where there are only two commodities at given prices, food baskets and fun baskets. We can also regard the quantity of money the agent has at his disposal as part of the situation. But these facts do not tell us in any way what the agent's rational purchasing activities in the situation will look like. In order to find out about that, we must also know something about the agent, for example that he is something of a hedonist and has well-defined indifference curves. One must be acquainted with the agent's preference system. If this has a certain character, then a certain way of acting is rational in the situation; if it has some other character, then some other way is rational. Two sorts of agents can find themselves in the same situation without the same sort of action being rational for both of them. Point two must be given the following formulation:

(2) *Analysis of the situation:*

In a situation of type *C*, the appropriate thing for an agent of type *A* to do is *d*.

An action is thus never rational in itself, but rational for a *certain sort* of agent in a *certain sort* of situation. But more needs to be said about the agent. What *relation* is there between the agent and the situation? Points (1), (2), and (3) above say nothing about this. Point (1) says, for example, only that (given the modification just introduced): 'Agent *a* of type *A* was in a situation of type *C*.' But the answer to the question about the relation of mediation between agent and situation is not hard to find. It is the intentionality category which supplies this relation – i.e. usually perceptions, but also other intentional acts and states. Point (1) must be rewritten in the following way:

(1) *Description of the situation:*

- (a) Agent *a* of type *A* was in a situation of type *C*.
- (b) Agent *a* was intentionally connected with situation *c* of type *C*.

It is not sufficient that *a* is in the situation *c*, he must also have

the situation as his intentional correlate. We can now describe the consumer example in the following way:

- (1) *Description of the situation:*
 - (a) Agent *a* of type *A* (hedonist with specified indifference curves) was in a situation *c* of type *C* (= found himself in a market with only food baskets and fun baskets at given prices, as well as having a sum of money to spend).
 - (b) Agent *a* correctly apprehended *c* as *C* and correctly apprehended himself as being an *A*. (Agent *a*: 'Given that I am a "foodie" and see a jar of pâté de foie gras in the food basket what shall I do?')
- (2) *Analysis of the situation:*
In a situation of type *C*, the appropriate thing for an agent of type *A* to do is *d* (= buy *x* food baskets and *y* fun baskets).
- (3) *Rationality principle:*
 - (a) Agents normally decide to act in a way appropriate to the situations in which they find themselves.
 - (b) Agents can be affected by non-rational factors.
- (4) *Closure clause:*
 - (a) *a* decided to act rationally
 - (b) *a* was not affected by non-rational factors in *c*.
- (5) *Explanandum:*
(Therefore) *a* did *d*.

In the structure given here, certain things are easily seen. The remarks made about irrational factors and about spontaneity lead to a completely new premise, point (4), and that in its turn makes point (3) superfluous. Thus we see by chance why actions and explanations of actions in many traditions are understood as unique and as more bound to individual situations than are explanations in the natural sciences. Point (3) is universal, but point (4) is determined by the situation. One can also see that *a*'s considerations (i.e. some of *a*'s relevant mental actions), which lead to his analysis of the situation, are not included as particular actions in the schema. They ought to be, but their omission does not affect the conclusion I wish to draw with the help of this schema.

The conclusion is that the agent concept does not require us to introduce any part-whole relations over and above those required

by the subject category. The intentionality category is already a defining characteristic of the subject, and the spontaneity category does not lead to any new part-whole relations.

Before we get involved in some of the part-whole relations based on intentionality, a few words should be said about misintentionality and contradictory intentionality.

15.3 MISINTENTIONALITY AND CONTRADICTORY INTENTIONALITY

In the example of the previous paragraph the satisfaction modality of intentional acts and states always had the value 'satisfied'. We shall now see what happens if we allow it to receive the values 'partially satisfied' and 'unsatisfied'. The interesting cases involve agents. Then we sometimes get what can be called the 'glass-door effect'. If we are walking along a corridor which we perceive correctly until we come to a glass door which we do not see at all, then we walk into the door. Such an action or occurrence can only be explained with the help of *misintentionality*, in the present case a misperception. We walked into the door because we did not see it. The explanation has a structure which links up with the structure of the position of methodological individualism, but which extends it as follows:

- (1) *Description of the situation:*
 - (a) Agent *a* of type *A* was in a situation *c* of type *C*.
 - (b) Agent *a* was intentionally connected with *c* as being of type *C*¹ and with himself as being of type *A*.
- (2) *Analysis of the situation:*
 - (a) In a situation of type *c*, the appropriate thing for an agent of type *A* to do is *d*.
 - (b) In a situation of type *C*¹, the appropriate thing for an agent or type *A* to do is *d*¹.
- (3) *Rationality principle:*
 - (a) Agents normally decide to act in a way appropriate to the situations in which they find themselves.
 - (b) Agents can be affected by irrational factors.
- (4) *Closure clause:*
 - (a) *a* decided to act rationally.
 - (b) *a* was not affected by non-rational factors in *c*.

- (5) *Explanandum*:
 (Therefore) *a tried* to do d^1 .

In the glass door case, C is to be equated with the corridor as it actually is, i.e. with a glass door. And the rational action (d) for the given person is to go through the corridor and open the glass door. C^1 , on the other hand, stands for the corridor without a glass door, and d^1 is the action of simply walking through the corridor. In the schema itself there is only the conclusion that a can try to walk straight through the corridor. But actually the schema contains more. C and d^1 can be compared. It is this comparison which leads to the conclusion that a is going to collide with the door. However, there need not always be a collision, literally or metaphorically, between an intended action and a misperceived situation. Let us look at the example of the consumer and assume that the agent misunderstands the prices, i.e. has an act that is characterized by misintentionality. He then buys z food baskets and u fun baskets, which is rational in a situation of type C^1 ; not x food baskets and y fun baskets, which would be rational in a situation of type C . A comparison of d^1 (the buying of z food baskets and u fun baskets) and C can be made, and its result is that d^1 is perfectly feasible. The conclusion here is not that d^1 is *impossible*, but that a performs a *non-optimal* act because he misunderstands the situation.

The comparison between d^1 and C is a comparison of a potential action with an actual situation. The type of universals which actions and naturalistic situations instantiate has been discussed in earlier chapters. The type of comparison that is made between d^1 and C is not a categorially new type of comparison. It leads to the same type of possibility or impossibility theorems as was discussed in the last chapter. In the same way as we can see that a ball with a 2-cm diameter cannot go through a hole with a 1-cm diameter, and can see that in a thermostat system with negative feedback the temperature cannot vary, so we can see in the glass door case that a collision is going to occur. In the consumer example we can see that a non-optimal action will be carried out.

The comparison between the real situation C and the situation which is the (not yet realized) intentional correlate is of course a comparison made by an external onlooker. The agent himself only acts, and then the collision just happens. The agent cannot make the comparison made by the external onlooker.

Presentational intentionality is in several senses primary, but representational intentionality is just as real. If we assert that at the same time both p and not- p is the case, where p is a naturalistic state of affairs, then we are asserting something contradictory. The assertion itself (the representational intentionality) occurs, but it is logically impossible for its intentional correlate to exist. Contradictions do not exist, but intentional acts directed to contradictory states of affairs do occur. When I now speak of contradictions I mean logical contradictions in the extended sense of logical I earlier (in section 8.2) described. Logical insights are not confined to insights which are based on formal-logical relations. There is material logic or existential dependence as well; one case is provided by the inclusion relations between inclusive determinates of one determinable. In this sense, one has a logically impossible intentional correlate if one believes that a 2-cm ball can go through a 1-cm hole, or that a volume of 5 cm³ includes a volume of 10-cm³. All of the impossibility theorems I spoke of in the previous chapter actually say that a certain intentional correlate is logically impossible.

I would thus like to extend the concept of a contradiction, or, more correctly, return it to approximately the position it occupied before modern formal logic replaced it by a very watered-down counterpart.⁵ On the other hand, I hold fast to the view that contradictions cannot exist in the non-intentional part of the world.

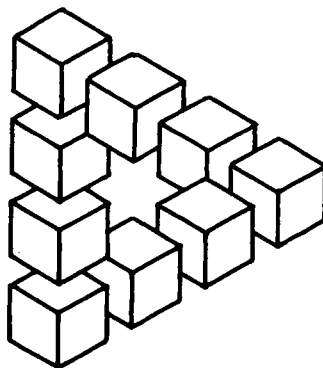


Figure 15.1
(Opus 1, 1934 Design Oscar Reutersvärd)

Since presentational intentionality claims to be in direct contact with the world, the conclusion is that *only representational intentionality* can be directed toward contradictions. But, as we shall now see, representational intentionality actually contains perception as well as thinking in the ordinary sense.

All of the impossible figures due to artists such as Escher and Reutersvärd (see Figure 15.1) are such that when we direct our attention to them our acts have contradictory correlates. Their intentional correlates are logically impossible. This point is easily misunderstood if presentational and representational intentionality are confused. Is the impossible figure the correlate of a presentational or of a representational intentional act? The solution of the problem lies in the insight that seeing a picture, like most uses of ordinary language, contains something which is, so to speak, presentationally given, without for that reason ceasing to be a representational act. The purely graphical pattern (like the purely linguistic sign) is *presentationally given* (i.e. it makes no claim to go beyond the act in question), but there is no presentational act *directed* to the graphical signs. To see a picture is to have a representational intentional act.

In an impossible figure different parts of the drawing itself point toward different states of affairs, states of affairs which are logically impossible to add to one another. In the same way different statements may simultaneously point toward states of affairs which it is logically impossible to combine. I believe, however, that picture-contradiction is genetically primary and that from a developmental perspective linguistic contradiction should be seen as secondary.

15.4 NESTED INTENTIONALITY

We shall now look at the interaction between two subjects or two agents. I would once again like to remark that our discussion takes place on the categorial plane. It is therefore very abstract. We shall clarify some different types of possible intentionality structures; we shall investigate whether there appear new such structures when an intentional act is directed toward another act. I shall approach the abstract relations we want to isolate by looking at Jean-Paul Sartre's analysis of the phenomenon of shame. Sartre asks his readers to imagine a situation where one makes a vulgar gesture:

I have just made an awkward or vulgar gesture. This gesture clings to me; I neither judge it nor blame it. I simply live it. I realize it in the mode of for-itself. But now suddenly I raise my head. Somebody was there and has seen me. Suddenly I realize the vulgarity of my gesture, and I am ashamed.⁶

We shall try to place this simple but keen observation into a more abstract structure. At least the following moments belong to the situation described:

- (1) the agents A and B , as well as the action (= the vulgar gesture of $A = Ad$);
- (2) the fact that A knows that he is performing action d , and the fact that B sees and despises A for doing d ;
- (3) the fact that A sees that B sees that Ad .

This structure can be made clear if we extend the notation used above and use an arrow to represent intentionality in general; specific kinds of intentionality are represented by an arrow with the specific character described above it. With such a notation we obtain the following:

- | | | | |
|-----|--|--|------------------------|
| | A | | B |
| (1) | Ad | | B |
| | knows | | despises |
| (2) | $A \longrightarrow Ad$ | | $B \longrightarrow Ad$ |
| | sees | | despises |
| (3) | $A \longrightarrow (B \longrightarrow Ad)$ | | |

One of Sartre's main conclusions can now be stated in the following way. Shame can never be found on plane (1) or (2), but requires the structure which exists on plane (3). Shame cannot be identified with the action Ad , and neither can it be identified with A 's being aware that he is doing d . Shame cannot exist in me as a monad, but presupposes two subjects in intentional contact with one another. Sartre himself gives a concise formulation of the points: 'Thus shame is shame *of oneself before the other*.'⁷ On the first plane there exist only states of affairs without any intentionality, and shame cannot be found there since shame essentially involves intentionality. On the second plane we find intentional acts directed towards non-intentional states of affairs, but shame cannot be found here either since shame implies that one is ashamed about something *before someone else*. Shame includes an intentional act

directed towards another intentionality. Shame is a type of whole which presupposes at least *two participants*.

In order to understand this analysis, certain things must once again be stressed regarding the relation between representational and presentational intentionality. Sartre's example falls under the category of presentational intentionality. However, this does not imply that one cannot feel ashamed when alone, but rather that, when one is ashamed and alone, the same fundamental structure obtains; the only difference being that person *B* is then not a real person but an imagined one, perhaps even an imagined undetermined 'Mr Somebody'. Sartre emphasizes this point and maintains not only that the same structure is to be found in presentational and representational shame, but also that representational shame is only to be found if it has actually been preceded by presentational shame:

Yet although certain complex forms derived from shame can appear on the reflective plane, shame is not originally a phenomenon of reflection. In fact no matter what results one can obtain in solitude by the religious *practice* of shame, it is in its primary structure shame *before somebody*.⁸

With this in mind, we can return to the abstract structure described above. It is not only the case that there exist three distinct planes, there also exist relations among the planes. What exists or occurs on plane (3) implies that which exists on planes (2) and (1). That $A \rightarrow (B \rightarrow Ad)$ entails *A*, *B*, and *Ad* as well as $(B \rightarrow Ad)$, ought to be clear without further comment; but it also entails $A \rightarrow Ad$. *A* cannot see that *B* sees that he does *d* without thereby becoming aware that he does *d*. This means that planes (1) and (2) are necessary conditions for $A \rightarrow (B \rightarrow Ad)$, or that the latter is one-sidedly existentially dependent on the former.

The existential dependence relation between planes is a dependence relation between *satisfied* presentational intentional acts. Where the acts are not satisfied the dependence relation obtains within the act of a single agent; the structure described is then not a structure between two real agents but merely a structure that connects this agent and an imagined agent. The naive realism I defended in section 13.6 allows me to regard real agents as normally nested.

When an intentional act of a subject is existentially dependent

upon intentional acts of other subjects, we have a very specific kind of ontological unity which I shall call *nested intentionality*. Such unities can be more or less nested, but the fundamental structure of nested intentionality is $A \rightarrow (B \rightarrow A)$. It should be compared with intentional phenomena where an agent is directed towards a naturalistic state of affairs ($= A \rightarrow O$, where O stands for object), or where an agent sees that another agent sees something; either a naturalistic state of affairs ($= A \rightarrow (B \rightarrow O)$) or a third agent ($=$

$A \rightarrow O$	}	
$A \rightarrow (B \rightarrow O)$		
$A \rightarrow (B \rightarrow C)$		
$A \rightarrow (B \rightarrow A)$		
		= non-nested intentionality
		= nested intentionality

The last structure above is an abstract intentional structure which is exemplified by the intentional moment of shame (shame of course involves more than intentionality; cf. page 262). But this nested structure is more abstract than the particular example we have described; there can exist other examples of the same structure. The structure in question delimits the intentional moment of a whole set of emotions and intentional states. Apart from shame, it also fits its opposite, pride. To be proud is primarily the same as being proud of oneself *before another*, to paraphrase Sartre. Pride, too, is a whole which has participants as parts.

Emotions like the opposites pride–shame thus distinguish themselves from emotional opposites like happiness–sadness, pleasure–horror, desire–repulsion, admire–despise, in that the latter pairs can (they need not) have non-nested intentional structures like $A \rightarrow O$ or $A \rightarrow B$ while the first pair necessarily involves the nested structure $A \rightarrow (B \rightarrow A)$. This also means that one kind of emotion may be existentially dependent on another kind of emotion. Shame depends upon despising someone and pride depends upon admiration:

$$\begin{array}{l}
 A \text{ is ashamed: } A \xrightarrow{\text{sees}} (B \xrightarrow{\text{despises}} A) \\
 A \text{ is proud: } A \xrightarrow{\text{sees}} (B \xrightarrow{\text{admires}} A)
 \end{array}$$

It is not only mental states which fit into such structures. A person cannot have the property of being a teacher in splendid isolation in the way that he *in himself* has a certain length and a

certain shape. Nor can he have the property of being a teacher by himself in the way in which he can perform actions like sitting, walking, drawing, and speaking. Assume, for example, that I day after day act just like a teacher. At some time every day I go to a certain place and give lectures. But nobody sits and listens, and I am not paid by any institution. In that case I am not a teacher in spite of my manifesting *precisely the same behaviour* as a teacher. If there had been students sitting there and had I been paid by a school or university the same behaviour would have made me a teacher. The property of being a teacher is, like shame and pride, dependent upon something beyond behaviour itself; it presupposes certain types of nested presentational intentionality. In order to be a teacher I must *see* that there exist students *who see me*. The teacher must see that there are students who see him as a teacher. We obtain the structure: $T \rightarrow (S \rightarrow T)$. Nor, in a corresponding way, can one be a student *per se*, but must fulfil the intentionality structure $S \rightarrow (T \rightarrow S)$.

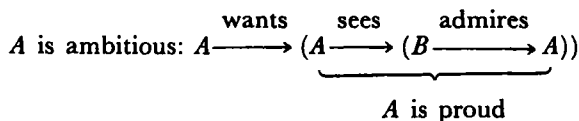
In chapter 8 the 'teacher-student' couple was used as an example of internal relations. Together with 'preacher-congregation' and 'capitalist-worker' it was the only example where there was a spatial gap between relata. We see now what this gap depends on. The existential dependence relations are in these cases based on the intentionality category. The relata are *participants*.

It might seem amazing that social roles or 'character masks', to borrow an expression of Marx, exemplify the same structure as emotional states. Social roles and emotional states seem otherwise to be categorially distinct. That this is so depends on the fact that we analyse only the intentionality moment of the phenomena; we have left aside the tendential moment with its accompanying actions. And it is on this latter side that the differences exist which we so clearly experience. Emotional states fall under the tendency category, while roles seem to fall under the spontaneity category. Of course roles are completely deterministically understood in many sociological role theories, i.e. they become tendencies, but in everyday speech associations which are connected with the spontaneity category (free will, etc.) seem to have the upper hand.

The structure we have now studied, $A \rightarrow (B \rightarrow A)$, can be easily built into similar but more comprehensive nested structures. Let us return to Sartre's example of the man who is ashamed. If, when he is ashamed, he begins to blush about his own shame, the observer

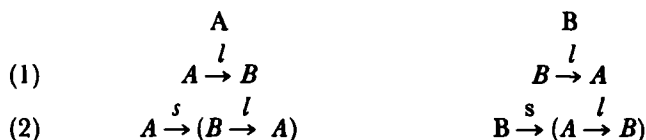
can also perceive *that* the man is ashamed. The man, in turn, can discover this and even begin to be ashamed of being ashamed. He thinks it contemptible to be ashamed. We have a shame of second order whose intentionality is nested according to the following structure: $A \rightarrow (B \rightarrow (A \rightarrow (B \rightarrow Ad)))$, i.e. A is ashamed because A sees that B despises A for being ashamed that B despises what A is doing.

Another complex nested structure is the intentional moment of being ambitious. A man who is ambitious wants or strives for situations where he can see that others (B) admire him. We get:



What has been said about pride, shame, shame of second order, and ambitions obviously shows the possibility of a formal study of emotions and other intentional states; a study based on the concept of nested intentionality. Descartes and Spinoza tried to carry out such analyses of the soul's passions and emotions, but since then it has not, to say the least, been *comme il faut* amongst philosophers, except for work by Brentano and his pupils.

All of the intentional states hitherto mentioned are states which via intentionality depend upon and contain other persons, but which, *as states*, exists in one particular person. Let us now investigate a phenomenon like friendship. That A and B are friends seems to mean more than that A likes B and that B likes A (row (1) below). There must be some sort of mutuality that goes deeper.⁹ In analogy with earlier analyses, we can perhaps say that A in addition must *see* that B likes him, and that B must *see* that A likes him (row (2) below). If we return to the abstract notation given earlier, and abbreviate 'likes' with ' l ' and 'sees' with ' s ', we can write the following with A and B as agents:



To my mind, not even this amendment is enough in order to capture the intentional moment of friendship. A must *want* B to see

that *A* likes him; and *B* must *want* *A* to see that *B* likes him. We get yet another row ('want' = 'w'):

$$(3) A \xrightarrow{w} (B \xrightarrow{s} (A \xrightarrow{l} B))$$

$$(= A3) \quad B \xrightarrow{w} (A \xrightarrow{s} (B \xrightarrow{l} A)) (= B3)$$

If there is real friendship, *B* should, I think, see *A*'s want as described by (A3) and *A* should see *B*'s want as described by (B3). Furthermore, *A* should want *B* to see this, and *B* should want *A* to see this. We have to add another two rows:

$$(4) A \xrightarrow{s} B3 (= A4) \qquad B \xrightarrow{s} A3 (= B4)$$

$$(5) A \xrightarrow{w} B4 \qquad B \xrightarrow{w} A4$$

When one knows that *p*, reflection shows that one also knows that one knows that *p*, as well as that one knows that one knows that one knows that *p*. There is no end to this regress, and in the same way the intentional moment of friendship involves an infinite series. Every other row gets '*A* sees ...' and '*B* sees ...', and every other row gets '*A* wants ...' and '*B* wants ...', where the intentional correlate is to be found in the opposite column on the row before. This regress has obvious similarities with the regress which occurs in two mirrors which reflect one another, I shall thus call it 'intentional mirror infinity', or, for short, 'intentional mirror'.

My thesis is quite simply that intentional mirrors constitute a special sort of infinity. They are of course *potential* infinities in the same sense as a spatial extension is potentially divisible to infinity. But spatial extension and intentionality are as different as categories can be, and so we must have two completely different kinds of infinity. That intentionality can involve infinity is neither more nor less disturbing than the role of infinity in other categories.

An analysis of the intentionality moment of phenomena of the mutual type – like friendship, love, and loyalty – must result in an infinite regress of the type described. Even if *A* likes *B* and knows that *B* likes him, friendship is more than this. It contains the possibility that *B* should know that *A* knows that *B* likes him, and the possibility that *A* should know this, and so on. Since the intentionality category is such that the subject-pole is anchored in only one body-substratum, there can be no question of any

absolute mutuality in the sense of the complete merging into one another that mystics claim to describe. The closest we can come to one another as subjects is to stand in the mutuality relation exhibited by intentional mirrors.

The fundamental intentional mirror on which all other social phenomena are at bottom founded, is the following (ϕ = 'perceive'):

$$\begin{array}{ll}
 & A \qquad \qquad \qquad B \\
 (1) & A \xrightarrow{\phi} B \qquad \qquad \qquad B \xrightarrow{\phi} A \\
 (2) & A \xrightarrow{\phi} (B \xrightarrow{\phi} A) \qquad \qquad \qquad B \xrightarrow{\phi} (A \xrightarrow{\phi} B) \\
 (3) & A \xrightarrow{\phi} (B \xrightarrow{\phi} (A \xrightarrow{\phi} B)) \qquad \qquad \qquad B \xrightarrow{\phi} (A \xrightarrow{\phi} (B \xrightarrow{\phi} A)) \\
 & \qquad \qquad \qquad \text{[etc.]}
 \end{array}$$

Let us abbreviate this structure by $A \leftrightarrow B$, and expand the example by bringing in a third agent, C . We thereby also get nestings of A and C as well as of B and C according to the same structure, i.e. $A \leftrightarrow C$ and $B \leftrightarrow C$. But there also obtains a new *kind* of structure based on the fact that C (we assume) perceives the nestedness $A \leftrightarrow B$, B perceives $A \leftrightarrow C$ and A perceives $B \leftrightarrow C$. On this basis another intentional mirror arises:

$$\begin{array}{lll}
 & A \qquad \qquad \qquad B \qquad \qquad \qquad C \\
 (1) & A \rightarrow (B \leftrightarrow C) \qquad \qquad \qquad B \rightarrow (A \leftrightarrow C) \qquad \qquad \qquad C \rightarrow (A \leftrightarrow B) \\
 (2) & A \rightarrow (B \rightarrow (A \leftrightarrow C)) \qquad \qquad \qquad B \rightarrow (A \rightarrow (B \leftrightarrow C)) \qquad \qquad \qquad C \rightarrow (A \rightarrow (B \leftrightarrow C)) \\
 & \quad A \rightarrow (C \rightarrow (A \leftrightarrow B)) \qquad \qquad \qquad B \rightarrow (C \rightarrow (A \leftrightarrow B)) \qquad \qquad \qquad C \rightarrow (B \rightarrow (A \leftrightarrow C)) \\
 (3) & A \rightarrow (B \rightarrow (A \rightarrow (B \leftrightarrow C))) \qquad \qquad \qquad \cdot \qquad \qquad \qquad \cdot \\
 & \quad A \rightarrow (B \rightarrow (C \rightarrow (A \leftrightarrow B))) \qquad \qquad \qquad \cdot \qquad \qquad \qquad \cdot \\
 & \quad A \rightarrow (C \rightarrow (A \rightarrow (B \leftrightarrow C))) \qquad \qquad \qquad \cdot \qquad \qquad \qquad \cdot \\
 & \quad A \rightarrow (C \rightarrow (B \rightarrow (A \leftrightarrow C))) \qquad \qquad \qquad \cdot \qquad \qquad \qquad \cdot \\
 & \qquad \qquad \qquad \text{[etc.]}
 \end{array}$$

The principle generating the successive higher order intentional levels is that A on level (2) takes as intentional correlates B 's and C 's intentionality on level (1), that A on level (3) takes as intentional correlates B 's and C 's intentionality on level (2), and so on; and in a corresponding way for B and C . Each level will in this way contain 2^{n-1} different intentionalities for each person, where n is the number of the level. This infinity we can abbreviate by

$$\overline{\qquad \qquad \qquad} \\
 A, B, C$$

Now we bring in a fourth agent, *D*. First, this person is nested in $D \leftrightarrow A$, $D \leftrightarrow B$ and $D \leftrightarrow C$. Secondly, he is nested in

$\longleftrightarrow \quad \longleftrightarrow \quad \longleftrightarrow$
 $D, A, B, D, B, C \text{ and } D, A, C$

but also, thirdly, nested in a regress beginning with

\longleftrightarrow
 $D \rightarrow (A, B, C)$

and ending in an infinity which we can symbolize by

\longleftrightarrow
 $A, B, C, D.$

This kind of nested intentionality gives us the real structure of inter-subjectivity.¹⁰

My introduction of the category of nested intentionality took its departure in Sartre's analyses of emotions. In order to bring out some other features of nested intentionality it is pertinent to begin with H.P. Grice's attempted analysis of meaning.¹¹ I say 'attempted' because damaging criticisms have been made of Grice's analyses, but such criticism should not (and usually has not tried to) conjure away the type of nested intentional phenomena described.¹² I am not here interested in the specific details of different speech acts, I am interested in showing the existence of different general structures of nested intentionality.

A common way of restating Grice's analysis very briefly is the following: *S* meant something by (or in) uttering *x* iff *S* uttered *x* intending

- (1) that his utterance of *x* produce a certain response *r* in a certain audience *A*;
- (2) that *A* recognize *S*'s intention (1);
- (3) that *A*'s recognition of *S*'s intention (1) shall function as at least part of *A*'s reason for *A*'s response *r*.¹³

As indicated above this analysis has flaws as a general analysis of meaning, but, in my opinion, it captures the phenomenon of giving an order very neatly.

Assume that *A* says to *B*: 'Open the window!', or via gestures transmits the same message. If *B* willingly follows the order, he performs an action the *reason* for which is that another person (*A*) wants him to do it. If *B* has understood the order, he has not only

understood that *A* wants him to open the window, but also that *A* wants it to be the case that *B* will realize that *A* wants him to open the window. *A* does not say that *B* should open the window for an arbitrary reason, but intends the order in itself to be a sufficient reason; so we must add that *A* wants *B* to open the window *because B realized what A wants him to realize*. The structure in the order fits the following schema: *A* orders *B* to do *d* if and only if

- (1) *A* intends *B* to do *d*;
- (2) *A* intends *B* to recognize (1);
- (3) *A* intends *B* to do *d* for the reason that *B* recognizes (2).

If we symbolize *B*'s doing of *d* with *Bd*, and the fact that *Bd* is carried out for a certain reason with *Bd* [*r*: . . .], we can capture the structure of obeying an order (an execution of the behaviour commanded) in a way which shows its resemblance to the other structures discussed above. Let '*i*' = 'intends' and '*r*' = 'recognizes' when '*r*' stands above the arrows).

<i>A</i>	<i>B</i>
(1) $A \xrightarrow{i} Bd$	$B \xrightarrow{r} (A \xrightarrow{i} Bd)$
(2) $A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd))$	$B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd)))$
(3) $A \xrightarrow{i} Bd$	Bd
$[r: B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd)))]$	$[r: B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd)))]$

The complex state of affairs described here is a whole which has *A* and *B* as parts in the sense of participants. The main differences between this nested intentionality and the structures described earlier are

- (a) that it contains reasons for actions;
- (b) the peculiar Gricean kind of self-reference which is to be found exemplified by *A* at level (3).

In 15.8 ('Macroagents') the importance of the latter will become clear.

15.5 NESTED MISINTENTIONALITY

I have earlier shown, in connection with the discussion of the

explanation schema of methodological individualism, how misintentionality in the form of a misperception can lead to non-optimal or unsuccessful actions ('the glass door effect'). The correlate of the intentional act was a naturalistic state of affairs. Let us now look briefly at the corresponding phenomena where we have two agents who stand to each other as intentional correlates. I shall now take my examples from a philosophically oriented psychologist partly inspired by Sartre, R.D. Laing.

Before I get involved in examples, it may be of interest to cite from Laing's foreword to his book *Knots*. He writes:

The patterns delineated here have not yet been classified by a Linnaeus of human bondage. They are all, perhaps, strangely, familiar. . . . I could have remained closer to the 'raw' data in which these patterns appear. I could have distilled them further towards an abstract logico-mathematical calculus.¹⁴

I should add that it is not my aim to emulate Linnaeus. On the other hand, I do want to bring out a little of the abstract calculus Laing also believes possible. By raising Laing's example to a higher level of abstraction, one sees similarities in structure to the analyses of Sartre and Grice, but also dissimilarities. Laing has actually captured a type of intentionality phenomenon which the others have not noticed. As we shall see yet again nested intentionality provides an area of study which is well worth philosophical cultivation. Now to the first example:

She wants him to want her	
He wants her to want him	
To get him to want her	
she pretends she wants him	
To get her to want him	
he pretends he wants her	
Jack wants	Jill wants
Jill's want of Jack	Jack's want of Jill
so	so
Jack tells Jill	Jill tells Jack
Jack wants Jill	Jill wants Jack
a perfect contract ¹⁵	

Jack's intentional correlate is 'Jill wants me', but he does not reach that correlate. The state of affairs which exists is rather 'Jill wants

Jack to want her'. In the same way and for the same reason Jill does not reach her correlate either. The two misintentionalities bind each other, however, in such a way that they all the while apparently confirm one another. Nested misintentionality will come to determine many of their actions. Laing calls such cases as these 'fantasy circles'. The circles are seldom seen through even by outside observers, which is, I think, the reason for our still lacking labels for this type of social-psychological phenomenon.

Behind Jack's (and Jill's) pretending there lies a form of strategic thinking which is connected to the schema of explanation in methodological individualism. Jack understands the situation as though it were rational for him to trick Jill in order to reach his correlate. His goal is that she will want to have him, and he analyses the situation as being such that *if* Jill believes that Jack wants to have Jill, *then* Jill will come to want to have Jack. But since Jack does not primarily want to have Jill but only Jill's *wanting-to-have-Jack*, Jack has to lie, i.e. create a misintentionality in Jill.

In the earlier examples of misintentionality in agents, an understanding or grasp of that intentionality was a condition for understanding certain unsuccessful actions. In these examples intentionality was directed toward naturalistic states of affairs, but now we have intentional acts directed towards intentionality and misintentionality which are themselves directed towards intentional acts. We obtain a situation where the actions can be described neither as unsuccessful nor as irrational. The actions cannot be called unsuccessful since in one sense Jack and Jill both obtain what they want; and the actions cannot be called irrational for there exists no knowledge which *A* and *B* had but did not use.¹⁶

The notation I have used in order to represent nested intentionality can be enlarged so that some features of nested misintentionality can also be represented. When we have misintentionality we shall write a dash (–) under the arrow which represents the intentional act in question. Laing's example then contains the following structure ('want' = '*w*' and 'sees' = '*s*'):

	<i>A</i>		<i>B</i>
(1)	$A \overset{w}{\rightarrow} (B \overset{w}{\rightarrow} A)$		$B \overset{w}{\rightarrow} (A \overset{w}{\rightarrow} B)$
(2)	$A \overset{s}{\rightarrow} (B \overset{w}{\rightarrow} A)$		$B \overset{s}{\rightarrow} (A \overset{w}{\rightarrow} B)$

- $$\begin{array}{ll}
 (3) & A \xrightarrow{w} (B \xrightarrow{s} (A \xrightarrow{w} B)) \quad B \xrightarrow{w} (A \xrightarrow{s} (B \xrightarrow{w} A)) \\
 (4) & A \xrightarrow{s} (B \xrightarrow{w} (A \xrightarrow{s} (B \xrightarrow{w} A))) \quad B \xrightarrow{s} (A \xrightarrow{w} (B \xrightarrow{s} (A \xrightarrow{w} B))) \\
 & \quad \quad \quad [\text{etc.}]
 \end{array}$$

The structure above should be compared with that below which represents the nestedness of two agents who both really want each other (cf. the analysis of friendship on pages 275f.).

- $$\begin{array}{ll}
 & A & B \\
 (1) & A \xrightarrow{w} B & B \xrightarrow{w} A \\
 (2) & A \xrightarrow{s} (B \xrightarrow{w} A) & B \xrightarrow{s} (A \xrightarrow{w} B) \\
 (3) & A \xrightarrow{w} (B \xrightarrow{s} (A \xrightarrow{w} B)) & B \xrightarrow{w} (A \xrightarrow{s} (B \xrightarrow{w} A)) \\
 (4) & A \xrightarrow{s} (B \xrightarrow{w} (A \xrightarrow{s} (B \xrightarrow{w} A))) & B \xrightarrow{s} (A \xrightarrow{w} (B \xrightarrow{s} (A \xrightarrow{w} B))) \\
 & \quad \quad \quad [\text{etc.}]
 \end{array}$$

Rows number (2), (3), and (4) contain \xrightarrow{s} instead of the first occurrence of \xrightarrow{s} in rows number 2, 3, 4. But since this difference is not visible to the agents themselves, the similarity between these corresponding levels can hide the important fact, namely the difference between rows 1 and 1.

Jack and Jill, consciously or half consciously, have the intention of tricking one another, but the same type of bad (or good?) interpretation-circle can arise even when the agents' acts exhibit misintentionality which is not at all based on any intention to mislead. What follows is an example which contains no hidden intentions, but which nevertheless literally leads to a vicious circle – a vicious circle which is actually a spiral. In the example a temporal development of nested misintentionalities is described involving 'mismatched interpretations, expectancies, experiences, attributions, and counter-attributions'. 'It', to continue the quotation, 'starts to whirl something like this':

Peter:

- (1) I am upset.
- (2) Paul is acting very calm and dispassionate.

Paul:

- (1) Peter is upset.
- (2) I'll try to help him by remaining calm and just listening.

- | | |
|--|--|
| <p>(3) If Paul cared about me and wanted to help he would get involved and show some emotion also.</p> <p>(4) Paul knows that this upsets me.</p> <p>(5) If Paul knows that his behaviour upsets me, he must be intending to hurt me.</p> <p>(6) He must be cruel, sadistic. Maybe he gets pleasure out of it, etc.¹⁷</p> | <p>(3) He is getting even more upset. I must be even more calm.</p> <p>(4) He is accusing me of hurting him.</p> <p>(5) I'm really trying to help.</p> <p>(6) He must be projecting.</p> |
|--|--|

15.6 NESTED CONTRADICTIONS

Earlier in this chapter (15.3) I made the point that not only linguistic phenomena but also perceptions may be contradictory. Now, after having introduced the idea that emotions contain as an essential part a moment of intentionality (15.4), I can add that even intentional states may be contradictory. Once again I shall take an example from Sartre. His analysis of love is a very illustrative example.¹⁸

Sartre says that when one loves someone, then one wants the loved one in two completely incompatible ways. *A* partly wants the loved one (*B*) to have freely chosen to love him, i.e. the loved one's love ought to fall under the *spontaneity category*. But in part and at the same time, *A* also wants the loved one *B* to love him as a result of a passion which she finds irresistible. In the latter case, *A* wants to function as an *efficient cause* for the loved one's response. But a phenomenon cannot at one and the same time arise both spontaneously and as the result of an efficient cause. Love is a contradiction, an impossible figure.

Our analysis does not make love out to be something which, so to speak, only occurs inside a person's head. The contradiction characteristic of love is to be found in one's very perceivings. I remind the reader that I maintained that perceptions are a form of presentational intentionality and that they contain *representational* intentionality. When we perceive a tree, for example, its front side

is *presented*, but its farside is *represented*. (The way the tree has just been – immediate past – and is about to be are also *represented*.) In the same way, causes of actions are often represented at the same time as the actions themselves are presented. Without coming into conflict with the thesis that contradictions can only be found in *representational* intentionality (cf. page 270) (not presentational), I can therefore maintain that we can have contradictory perceptions. The situation is the same as in the case of the impossible figures. There a graphic pattern was presentationally given and a representational intentional act was directed towards two incompatible figures. In the same way love, too, can be an impossible figure. In the perception of the beloved one's behaviour there is a representational intentionality which is contradictory.

A contradiction between spontaneity and efficient causality manifests itself in many human contexts. It is in fact the contradiction which, in the present theory of categories, represents the old conflict between natural-causality and freedom-causality, to borrow terms from Kant.¹⁹

An agent who has contradictory intentional acts strives after the impossible. The correlate cannot be reached, but his intentional states and acts nevertheless affect his actions. The exact way in which this expresses itself cannot be understood from intentionality alone. It requires studies of each individual case, even if certain clues can be obtained from the contradictory correlate itself. Someone who believes that square circles can exist may perhaps be led to draw different figures for the rest of his life in a vain search for a satisfying representation. And he who seeks love in the contradictory form described by Sartre will perhaps spend his life going from one woman to the next, or spend his whole life in dissatisfied nagging at the woman he nevertheless cannot bring himself to leave (as the poet writes: 'We can die by it, if not live by love').²⁰ But one cannot determine in advance how an impossible intention will try to realize itself in detail. It might be useful here to mention Gregory Bateson's famous studies of the 'double-bind' phenomenon.²¹ I shall take a simple example. A mother who, after a long absence, welcomes her child by saying, 'Oh, finally. Oh but I love you!', but who with her whole body shows abhorrence and dislike towards her child, gives the child a contradictory message. The child perceives a contradiction. And it can, as Bateson has shown, have both long-lasting and devastating consequences if it is

a frequent occurrence. Contradictions do not only exist in the sphere of the intellect and logic, but also in the sphere of feeling and perception. However, contradictions do not appear outside the sphere of intentionality.

The Bateson example conforms to the following structure ('sees' = 's' and 'loves' = 'l'):

$$A \xrightarrow{s} (B \xrightarrow{l} A \ \& \ B \xrightarrow{-l} A)$$

Sartre's contradictory love has a structure which is somewhat more complex since it involves reasons, and causes behind the state of love. But if we build the latter into the symbolism – we let ' l_s ' mean 'loves spontaneously', and ' l_p ' loves because of an overwhelming passion – we obtain a similar structure since l_s and l_p exclude one another.²² We get:

$$A \xrightarrow{s} (B \xrightarrow{l_s} A \ \& \ B \xrightarrow{l_p} A)$$

In both cases we have nested intentionality and we have contradictory intentionality, but neither case exemplifies what I want to call *nested contradiction*. Nested contradiction is a special kind of contradictory intentionality which, unlike the two examples above, does not have the form $A \rightarrow [p \ \& \ \text{not-}p]$. In nested contradictions the contradiction itself is nested. The structure can be found by means of substitution in the Gricean analysis of ordering earlier presented (see pages 278–9). When A orders B to do d the following intentional structure was said to obtain where 'intends' = ' i ', 'recognizes' = ' r ' when above the arrows, and ' B does d for reasons r ' = ' Bd [r : . . .]':

$$\begin{array}{l}
 \begin{array}{c} A \\ (1) \quad A \xrightarrow{i} Bd \\ (2) \quad A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd)) \\ (3) \quad A \xrightarrow{i} Bd \ [r.B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd)))] \end{array} \\
 \begin{array}{c} B \\ (1) \quad B \xrightarrow{r} (A \xrightarrow{i} Bd) \\ (2) \quad B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd))) \end{array}
 \end{array}$$

$$(3) \quad Bd \ [r:B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} Bd)))]$$

In this structure we shall now try to insert the specific command 'Forget C!'. Or, to avoid possible misinterpretations, *A* orders *B*: 'You must forget C!'. It is thus a matter not only of forgetting, but of forgetting because of *A*'s authority. To have forgotten something implies that one lacks intentionality toward that thing. If we symbolize the lack of intentionality with an arrow with a minus sign above it ($\xrightarrow{-}$) we can exchange '*Bd*' for ' $B \xrightarrow{-} C$ ' in the structure above. We then get this structure:

$$\begin{aligned} & A \\ (1) \quad & A \xrightarrow{i} (B \xrightarrow{-} C) \\ (2) \quad & A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{-} C))) \\ (3) \quad & A \xrightarrow{i} (B \xrightarrow{-} C) [r:B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{-} C)))] \end{aligned}$$

$$\begin{aligned} & B \\ (1) \quad & B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{-} C)) \\ (2) \quad & B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{-} C)))) \\ (3) \quad & (B \xrightarrow{-} C) [r:B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{r} (A \xrightarrow{i} (B \xrightarrow{-} C)))] \end{aligned}$$

There are no contradictions to be found on the first two levels, but the third level exemplifies a contradiction. First, there is a contradiction since *A* requires that *B* shall be both aware and not aware of *C*; *B* is supposed to *remember to forget C*. Second, the contradiction does not appear in the form $A \rightarrow (p \ \& \ \text{not-}p)$, i.e. *A* does not give *B* the order 'Remember to forget C!'. This means that we have a contradiction but not an ordinary contradiction of the type explored by formal logic; we have a *nested contradiction*.

If agent *B* does not realize the contradictory character of the command he will strive after the impossible. He will be like the man who tries to draw a square circle.

Another example of the same structure is the order 'Be spontaneous!'; to be spontaneous here means carrying out actions without bothering about the reasons for these actions. If *A* directs such an order at *B* we should exchange

$$A \xrightarrow{i} (B \xrightarrow{-} C)$$

on the first level in the last schema for

$$A \xrightarrow{i} (B \xrightarrow{-} r)$$

The latter means that *A* intends that *B* should not think at all of giving reasons. On the third level we then get a nested contradiction.

Nested contradictions are perhaps the grain of truth contained in Hegel's famous master-slave dialectic. The master (*A*) wants (1) the slave (*B*) to regard himself as a thing, i.e. as something that lacks intentionality:

$$A \xrightarrow{w} (B \xrightarrow{-} x)$$

A also wants *B* to (2) recognize this intention and (3) to turn himself into an intentionless thing because of this recognition. If

$$A \xrightarrow{w} (B \xrightarrow{-} x)$$

is inserted for

$$A \rightarrow (B \xrightarrow{-} C)$$

in the last schema, we get the corresponding nested contradiction on the third level.

I shall end this section by quoting yet another 'knot' from Laing. Think about this:

A son should respect his father
 He should not have to be taught to respect his father
 It is something that is natural
 That's how I've brought up my son anyway

 Of course a father must be worthy of respect
 He can forfeit a son's respect
 But I hope at least that my son will respect me, if
 only for leaving him free to respect me or not²³

15.7 MARXIST DIALECTICS AND INTENTIONALITY

This theory of categories of mine is marxist in inspiration and (to a lesser extent) as regards the problems dealt with, but somewhat analytic in its way of handling these problems.²⁴ The resulting ontology, however, is most closely related to the positions of two philosophers who belong neither to marxism nor to analytic

philosophy: Aristotle and Husserl. In this section I shall comment on a problem within marxism, the problem of dialectic. This problem is discussed by a philosopher whose views do not have much in common with this theory of categories: Hegel. It was Hegel who put forward the view that *in the world* there exist actual *contradictions*, and that it is a dialectic between contradictory opposites which explains the world and its history.

Both the idea of nested intentionality and the idea of internal relations are to be found in Hegel. But the account given in chapter 9 of internal relations shows that the latter are a special case of Husserl's existential dependence relations. These are all passive, not productive in Hegel's sense. Moreover, Husserl allows for one-sided dependence relations. And, last but not least, they have nothing to do with contradiction in Hegel's sense.

Jon Elster, whose work is analytic in orientation, has suggested that there exist interesting observations in what Hegel and Marx say about dialectics and contradictions, but that what is valuable must be separated from its mystifying packaging.²⁵ In this way it can be placed into a traditional analytical framework where all types of contradiction are defined in terms of logical relations between sentences. Elster's analyses are very valuable and I shall take my point of departure from them, even though my conclusions diverge somewhat from his.

Elster's 'vindication of dialectics' consists of a denial of the two theses: (1) 'There are contradictions in reality', (2) 'An adequate description of reality must contain self-contradictory propositions'; but a defence of a third thesis : (3) 'There are situations in reality that can only be described by means of the concept of a logical contradiction.'²⁶ He also makes a distinction between 'contradictions of the mind' and 'contradictions of society', and in the former group between 'contradictory desires' and 'contradictory beliefs'. What I have said so far about 'contradictory intentionality' is, as can be seen from the footnotes, partly inspired by Elster's analyses. One such case is my reference to Sartre's analysis of love, an analysis Elster cites as an example of contradictory desire. He who loves in the way described by Sartre finds himself, according to Elster, in a real situation 'that can only be described by means of the concept of a logical contradiction'.

One respect in which Elster and I differ is that I regard the concepts of intentionality and existential exclusion (D9.5) as the

basic concepts needed in order to understand contradictions, while his basic concepts are taken from formal logic. This difference is often of no significance, since logical contradictions belong to the area of the intentional (cf. sections 15.3 and 15.6); but significant differences sometimes show themselves. One such is my emphasis on contradictory perceptions. Perceptions are intentional phenomena, and, as such, they can in principle include contradictions. But for Elster, who begins with *descriptions*, all contradictions come automatically to be placed on the linguistic level; this is implied by thesis (3) above.

The distinction between 'contradictions of the mind' and 'contradictions of society' is not essential. What is essential is rather to show that classical examples of supposedly dialectical phenomena can be understood with the help of the intentionality category. For reasons of space I shall here take up only a couple of marxist examples. Certain supposedly dialectical phenomena can, as Elster has shown, be understood by means of 'non-dialectical' concepts like 'counterfinality' and 'suboptimality'. Both concepts belong to the intentionality category. I shall not try to summarize those of his arguments which relate to these concepts (since I agree with what he says), but shall comment on what he calls the 'fallacy of composition'²⁷ and discuss an example he does not deal with at all.

'The fallacy of composition' is the mistaken belief that because something is possible for each entity of a certain kind, then it must also be possible for *all* of them *simultaneously*. If we use formal notation this means that the following is not a logical truth: $(\forall x)(P(Fx) \rightarrow P(Vx) (Fx))$. It is easy to find examples where the antecedent is true but the consequent false. It is possible for each person who starts in a marathon race to win it, but it is not possible for all to do it together. There is only one winner, but any of those who start can be the winner. It is possible to read any of the books in my library in one day, but it is not possible to read all of them in one day. Some such examples may be called learning or reproduction impossibilities. It is possible for anyone at any time to break his promises, but it is not possible for everyone always to break their promises, for then the 'institution of promising' would not be able to reproduce itself, and no one would be able to learn what a promise is. A similar point can be made about the statement that not everyone can lie, for if that were the case, no one would be able to learn a language.

Elster's point is that a fallacious conclusion of this sort, which is a purely logical mistake by a certain person, can appear in a societal form. It is then not a question of a logical fallacy by a certain person, but rather a question of a societal situation which 'can only be described by means of the concept of (formal-)logical contradiction'. One of Elster's examples of this type of situation is what Marx calls 'the contradiction in the general formula of capital'. Assume that we have a society containing a group of producers and a group of merchants who make a profit by buying cheap and selling dear. Here it is possible both for each merchant and for all the merchants simultaneously to make a profit in this way. These merchant capitalists, whom we can assume to have become very well-off, might easily imagine the 'good society' to be a society consisting *only* of merchant capitalists. As a matter of fact a number of early economists made suggestions in this direction. Such a state of affairs is, however, impossible. Here we have a traditional fallacy of composition. It is possible for each merchant capitalist to buy cheap and sell dear, but it is not possible for everyone to buy and sell goods and to be merchant capitalists and live on such a profit-margin. Someone somewhere must lose what the other gains. So long as merchant capitalists only constitute a *part* of a society their profit can arise by being taken from those who are not merchant capitalists. But a society consisting only of merchant capitalists is a contradiction.

One should not, however, think that in the situation assumed above, the merchant capitalists themselves, or anyone at all, commit the fallacy of composition and actually believes that everyone in the society could be merchant capitalists. Each person can believe that he himself can live by buying cheap and selling dear, and not think at all about what is possible for society as a whole. Then there is no contradiction in the relevant intentional acts of any single person. There are no 'contradictions of the mind', but the *description* of the society includes a contradiction in the sense that a description of the aggregate of individual intentions cannot possibly be realized.

If, in our imaginary society, those who already are merchant capitalists strive to retain their status, and everyone else strives to acquire this status, then the contradictory 'intentional aggregate' leads to the occurrence somewhere of *conflicting tendencies*. We see here how the categories of contradictory intentionality and

counteracting tendencies are both exemplified. The social contradiction expresses itself in the fact that *A* wants to buy *B*'s goods below their value, while *B* wants to buy *A*'s goods below *their* value. There are two conflicting tendencies, and presumably there will be no exchange. The tendencies cancel one another. It is only this connection between contradictions and tendencies which makes Elster's thesis interesting. As I see it, Elster does not himself make this sufficiently clear. If such a connection does not exist, the thesis that 'There are situations in reality that can only be described by means of the concept of a logical contradiction' is reduced to the triviality that if one person believes *p*, and another not-*p*, then the whole situation in which they partake cannot be described in any other way than by means of a formal-logical contradiction. If one is to understand all the peculiarities of 'dialectics' more is needed than a distinction between contradictions on the one hand and conflicting tendencies on the other. One must also see the connection between them, i.e. see that intentions have both an intentional and a tendential aspect. Were this not the case, Elster's analysis would lack interest.

Marx himself speaks of the general formula of capital as $M-C-M^1$, where *C* stands for a commodity, *M* for a sum of money, and M^1 for a sum of money greater than *M*.²⁸ The general form represents the 'intention of capital', which is also the intentions of individual capitalists. One invests money (*M*) in buying commodities (*C*) in the hope of thereby obtaining even more money (M^1). Assume that all the individual capitalists are agents with this intention. But assume also that all these capitalists are merchant capitalists. Then yet another aspect of the 'contradiction in the general formula of capital' reveals itself. Since the formula $M-C-M^1$ is valid for each individual, but is not generalizable to all simultaneously, there must somewhere exist a limit to the number who can live with this formula as their guiding star. If the merchant capitalists have a conscious tendency to transform society in such a way that it consists only of merchant capitalists then they thereby unconsciously manifest a tendency to cease to be merchant capitalists. Marx can now be interpreted as saying that the formula $M-C-M^1$ also has a contradiction-free form. Industrial capital invests money (*M*), as does merchant capital, in the buying of goods (*C*), with the aim of obtaining more money (M^1). But here, according to Marx, there is no 'contradiction'. One way of realizing

the aim $M-C-M^1$ is to be a merchant capitalist, *another* is to be an industrial capitalist. If the first way contains an 'Elster-contradiction', then a *tendency* presumably arises to realize the same aim in another way, i.e. the *merchant* capitalists remain *capitalists* but change over to being *industrial* capitalists. This point must be borne in mind if one is to understand the whole of the 'dialectic' in Marx's example.

Elster picks out three examples of 'contradictions' in Marx's writings. I have now commented on one of these. The two others are 'the contradiction of the law of the falling rate of profit' and 'the contradiction between the forces of production and the relations of production'. With regard to these, however, I shall only refer to Elster's analyses,²⁹ and instead go through the example which Elster, remarkably enough, does not mention: Marx's famous 'derivation' of money in the first part of the first volume of *Capital*. It is here that one can see, more clearly than anywhere else, Marx's need of the category of existential dependence and of a well-delimited intentionality category.

Marx distinguishes among four forms of value which, from a chronological point of view, have followed upon one another in history, but which also in some never really clarified Hegelian (or Hegel-turned-upside-down) way, have been assumed to follow logically from one another.³⁰ The first form is the *simple* form of value. Historically speaking this form is connected with societies where the production of goods is only temporary and where trade is by direct barter. Exchange according to the *simple form* of value can be written in the following way:

$$x \text{ units of commodity } A = y \text{ units of commodity } B$$

The person who wants to sell A and buy B has no use for more A s but is in need of B s. Commodities of type B have, in Marx' terminology, *use value* for that person. The situation is correspondingly similar for the other person. Say that the two persons exchange food for clothing – the one is a farmer and the other a tailor. That the goods are exchanged in just the proportions given means that x units of A have the same *exchange value* as y units of B .

The second form is the *expanded form*, which can be written in the following way:

$$\begin{array}{l} y \text{ units of commodity } B \\ x \text{ units of commodity } A = z \text{ units of commodity } C \\ u \text{ units of commodity } D \end{array}$$

Here there are a number of different sorts of goods such that he who wants to sell *A* can immediately set *A* in relation to many other types of commodity. The exchange value can as a consequence be measured with the help of many alternative types of commodity. The third form, which Marx calls the *general value form*, is merely the reversal of the *expanded value form* (i.e. the second form); thus it has the following form:

$$\begin{array}{l} x \text{ units of commodity } A \\ y \text{ units of commodity } B \\ z \text{ units of commodity } C \\ u \text{ units of commodity } D \end{array} = v \text{ units of commodity } E$$

The last schema is typical for societies where the exchange of goods has an even greater intensity. The normal procedure when someone now wants to exchange commodity *A* for commodity *B* is that he first exchanges *A* for the *universal* equivalent *E*, and then with the help of *E* obtains *B*. Now the exchange value of all commodities are measured in one and the same commodity, *E*. In all three commodity forms all commodities have a use value for *someone*; this also holds for *E*. In the last case one can think of societies where a certain kind of cattle is used as currency. In the fourth and last form of commodity, the *form of money*, this relation is done away with. We have:

$$\begin{array}{l} x \text{ units of commodity } A \\ y \text{ units of commodity } B \\ z \text{ units of commodity } C \\ u \text{ units of commodity } D \end{array} = v \text{ units of money}$$

Money does not have use value in the normal sense, i.e. normally one cannot use money *directly* to satisfy human needs. It is not edible, it cannot be used as clothing, and so on. If one wants to speak of the use value of money, one might somewhat paradoxically say that its use value is to be the opposite of use value, namely, exchange value. Hegelian marxists have tried to read Hegelian conceptual development into the above four schemata. They begin with a supposed

conflict/contradiction between two properties, use value and exchange value. (In my opinion these two kinds of value are quite simply two mutually dependent moments of each commodity; but Hegelians see them not only as distinct but as contradictory.) This conflict/contradiction is then assumed to be 'transcended' (or *aufgehoben*) in that the one side of the conflict, the exchange value, necessarily releases itself from the commodity and takes on an existence of its own, namely, as money. The appearance of money is to be understood on this Hegelian interpretation as arising not from situationally determined intentions with accompanying non-intended consequences, but as a necessary consequence of a contradiction between use value and exchange value which inheres in commodities.

This, however, is Hegelian nonsense. The more developed forms do not follow logically from the less developed. But their less developed forms are conditions of existence for the developed form; the latter contain the former. In the states of affairs represented by the schema for the form of money, there necessarily exist the states of affairs represented by the schema for the general form, and so on. At the end of this chain of one-sided dependence there are the states of affairs which correspond to barter. These one-sided existential dependence relations are wrongly interpreted by Hegelians as a developmental schema (cf. page 133).

Money, however, cannot be completely ontologically understood in terms of these existential dependence relations. Special states of affairs connected with the intentionality category also play an important role. Marx himself strongly emphasizes that neither use value nor exchange value are properties inhering in things, but are relations. What *kind* of relations he has in mind he never makes clear; this lack I shall now try to remedy.

No commodity has use value in itself; use value is always use value for someone. One has to make a distinction between *direct* and *indirect* use value which Marx himself never explicitly made. If a commodity has direct use value, it can be used by the buyer in order to satisfy (in and of itself) some of the buyer's need(s). But if a commodity has indirect use value then it *cannot* in and of itself satisfy the buyer's need(s), but may well in another situation be exchanged for something which has direct use value. Direct use value might possibly be understood as a form of causal relation (i.e. the commodity satisfies a desire), but indirect use value can *only* be understood in terms of intentionality. If a commodity has

indirect use value for someone, then that someone understands that the goods are exchanged with the aim of making yet another exchange.

Let us look at some examples. We call the buyers and sellers P_1 , P_2 , etc. An exchange according to the simple value form can then be characterized in the following way (cf. the schema for order-giving on page 279):

- (1) P_1 knows that B has direct use value for him, and believes that A has direct use value for P_2 .
- (2) P_1 wants P_2 to recognize point (1).
- (3) P_1 wants P_2 to exchange B for A because P_2 recognizes point (2).

If the general value form is to fit into a similar schema, such a schema must look a bit different. We assume now that P_1 is in possession of the universal equivalent E , and wants to have commodity B :

- (1) P_1 knows
 - (a) that B has direct use value for P_1 ;
 - (b) that E has indirect use value for P_2 ;
 - (c) that E has direct use value for P_3 .
- (2) P_1 wants P_2 to recognize point (1).
- (3) P_1 wants P_2 to exchange B for E because P_2 recognizes point (2).

The buyers and sellers in these examples are ideal-typical in the normative sense. They do not intend to trick anyone, but want everyone to receive a use value at the right price. If we continue with this assumption, ordinary business with money (M) can be said to follow the schema below:

- (1) P_1 knows
 - (a) that B has direct use value for P_1 ;
 - (b) that M has indirect use value for P_2 ;
 - (c) that some commodity A, B, C, D, \dots , has direct use value for P_2 ;
 - (d) that M has indirect use value for P_x who sells the goods which have direct use value for P_2 ;
 - (e) that some commodity A, B, C, D, \dots , has direct use value for P_x .
- .
- .
- .

- (2) P_1 wants P_2 to recognize point (1).
- (3) P_1 wants P_2 to exchange B for M because P_2 recognizes point (2).

If P_1 is to be the *ideal-typical* merchant we have assumed him to be, he must have intentionality directed at an *infinite progression*. Since money does not have direct use value, it must be exchanged for something. He who has exchanged goods for money and cannot use the money has basically given his goods away without compensation. In reality, this need not be a question of an *infinite* progress, but only of a *progression into an unknown future*. But this is quite sufficient for the point I wish to make. Money is, like both language and machines, a human creation, but money is, from a categorial point of view, distinct from both of these in an interesting way.³¹

Money is distinct from machines in that the potentially endless interaction which is inherent in money can only be understood via the intentionality category. One can easily imagine machines (or natural processes, for that matter) which, given a certain supply of energy, consist of an eternal repetitive process, for example a machine which endlessly moves circular metal plates from place to place. But these discs cannot be money, for the machine cannot have intentionality which is required in order for indirect use value to exist. That which is specific to money is not a potentially endless interaction, but an *intentional* endless progression. Without intentionality currency notes are only paper, just as a letter without intentionality is only a pattern of ink strokes.

Money is distinct from language in that the material substratum plays a more important role. Money is a material part of a material exchange in a completely different way than the material substrata of linguistic signs are parts of linguistic interactions. Such substrata have no communication value in a communication. In the case of oral communication the substratum disappears more or less simultaneously with the arrival of the message. In the case of a written message, of course, the substratum remains, but the message itself can be read off by anyone who knows the language without that person's owning or standing in any intimate relation to the substratum. But in the case of buying and selling, as distinct from linguistic interaction, it is very important that the substratum is not dissipated and that it belongs to only one party in the interaction at a time.

Money is, it is worth repeating even today, a more noteworthy invention than most philosophers have realized.

15.8 MACROAGENTS

I began this chapter with a few words about atomism and holism. It is now time to deal with a specific part of this general issue, namely the difference between microagents and macroagents. By 'microagents' I mean individual persons, and by 'macroagents' such entities as organizations, companies, classes, and nations. When one nation declares war on another and when a company raises the price of its goods, it is macroagents who are acting. But these macroagents have, in a way, microagents as parts i.e. participants. How do microagents constitute macroagents?

The intentionality category is, in a sense, always connected to individual persons. There is no primitive macrointentionality. The subject-poles never unite into a single macrosubject-*pole* directed towards a certain intentional goal. A microagent is always founded on the same substratum as a subject pole; it is the spontaneity category which distinguishes agents from subjects in general. The actions of microagents can of course be aggregated in the traditional sense. If 10 individuals bake 3 loaves of bread each, then the 'unit' consisting of these 10 individuals seen as an aggregate, can be seen as having baked 30 loaves of bread. But this does not make the bakers into a macroagent.

When we now seek an understanding of what unites microagents to make a macroagent, we seek an understanding which refers to the macroagents' intentionality. On the pure action plane, as it manifests itself via the type of four-dimensional universals which was discussed in chapter 14, there is nothing special which distinguishes macroagents from patterns of a purely naturalistic type. Without intentionality, macroagents are reduced to a type of machine or organism. In the case of machines and organisms the parts add themselves to a whole with the help of spatio-temporal inclusive universals and the category of efficient causality. Macroagents, on the other hand, appear when a group of microagents has specific nested intentionality structures. As I have already indicated, a macroagent cannot have a subject-pole of its own which is distinct from the microagents' subject-poles. A macroagent cannot be defined in any other way than via the

specific nested intentionality structures of the constitutive micro-agents.

There are many different types of macroagent. I shall only bring out the structure of a few important cases which have a prominent place in political philosophy. Not in order to bring out everything that is involved in the cases at hand, but in order to show the importance of the category of nested intentionality. By means of this category one can explain the real difference between a society ruled by 'the general will' and a society ruled by 'the will of all'.

The last distinction is due to Jean-Jacques Rousseau. Rousseau has often been accused of being an unclear thinker, and many have complained about unclarities in the concept of 'general will' and its relation to the 'will of all'. I shall not try to refute these partly correct accusations, but I shall show that two specific intentionality patterns can be connected in a natural way with the two concepts. Neither of these patterns, I shall also show, fits an absolutely totalitarian society, a society ruled by a 'Leviathan' in Thomas Hobbes's sense. However, it will also be shown that there are an astonishing number of similarities between the nested structures of 'the general will' and absolute authority. But first let us look at Rousseau's distinction.

In what follows I shall only discuss cases where the macroagent in question has but two participants (= microagents). The generalization to more participants is in principle simple but cumbersome to write out, and I have earlier sketched the way such a generalization can be made (see pages 277–8).

The intentionality structure of the general will is closely related to that of friendship discussed in 15.4. Assume that the microagents A and B both want B to do a certain thing d ($= Bd$), and that they want this because they are part of a macroagent with a general will. Then the following obtains (w = 'wants'):

A	B
$A \xrightarrow{w} Bd$	$B \xrightarrow{w} Bd$
$A \xrightarrow{w} (B \xrightarrow{w} Bd)$	$B \xrightarrow{w} (A \xrightarrow{w} Bd)$
$A \xrightarrow{w} (B \xrightarrow{w} (A \xrightarrow{w} Bd \rightarrow))$	$B \xrightarrow{w} (A \xrightarrow{w} (B \xrightarrow{w} Bd))$
[etc.]	

On level (2) A wants B to want *what* A wants on level (1); and

vice versa for *B*. Furthermore, on level (3), *A* wants *B* to want that *A* wants what he wants on level (1); and vice versa for *B*. This should be compared with the following, the structure of the will of all.

	<i>A</i>	<i>B</i>
(1)	$A \xrightarrow{w} Bd$	$B \xrightarrow{w} Bd$
(2)	$A \xrightarrow{w} (A \xrightarrow{w} Bd)$	$B \xrightarrow{w} (B \xrightarrow{w} Bd)$
(3)	$A \xrightarrow{w} (A \xrightarrow{w} (A \xrightarrow{w} Bd \rightarrow))$	$B \xrightarrow{w} (B \xrightarrow{w} (B \xrightarrow{w} Bd))$
	[etc.]	

In a structure like this *A*'s want never becomes nested with *B*'s want. *A* wants what he wants and *B* wants what he wants, and there is nothing more to say. I have chosen an example where both microagents want the same, they both think that *B* ought to do *d*. If their wants had been different they would have had to construct or rely on some mechanism, for example democratic voting, in order to get a want for the macroagent.

The 'will of all' has a rock-bottom, or rather two, the wants of *A* and *B* on level (1). But the corresponding wants of the general will (level (1)) are not such rock-bottom wants, since the will of *A* is also supposed to conform to that of *B* and vice versa. There is no individual starting point, and that is the characteristic peculiarity of the general will. It ought to be noted, however, that such a will need not be static. The circle that *A* wants what *B* wants, who wants what *A* wants, requires only that *A*'s and *B*'s wants change *simultaneously* in the same way.

It ought also to be pointed out that the structure of the general will presumably never occurs in a pure form. The structure is to this extent an ideal type. There exist, however, it seems to me, many real groups which have a half-conscious striving towards the structure of the general will, or would at least like to appear outwardly to have this structure. In such a group conflicts are not allowed, and no one can really distinguish his individual will from the will of the others. In some way there just arises a general will for the group.

I now turn to the kind of macroagents where one microagent has absolute power and the other participants blindly follow. The latter have internalized the leader's authority in such a way that they

want him to decide what they shall want to do. If A is the ruler and B the ruled, and A wants B to do d , then similarly B wants B to do d . This is level (1) in the structure for *the general will*. Let us therefore see whether, on the higher-order levels, we can discern the difference which obviously must exist between the general will and an absolute authority.

On the second level in the general will, A wants B to want Bd , but that is also what our totalitarian leader wants, because he wants B to *want* to do what he is ordered to do. Similar arguments apply to the third level. In order to make apparent the difference we are interested in, we have to bring in the *reasons* A and B have for their wants. We shall once again use the notation $[r: \dots]$, only this time for symbolizing reasons not for an action but for a want.³² Taking such reasons into account, the general will looks as follows:

- A
- (1) $(A \xrightarrow{w} Bd) [r: B \xrightarrow{w} Bd]$
 - (2) $A \xrightarrow{w} (B \xrightarrow{w} Bd) [r: A \xrightarrow{w} Bd]$
 - (3) $A \xrightarrow{w} (B \xrightarrow{w} (A \xrightarrow{w} Bd)) [r: B \xrightarrow{w} Bd]$
[etc.]
- B
- (1) $(B \xrightarrow{w} Bd) [r: A \xrightarrow{w} Bd]$
 - (2) $B \xrightarrow{w} (A \xrightarrow{w} Bd) [r: B \xrightarrow{w} Bd]$
 - (3) $B \xrightarrow{w} (A \xrightarrow{w} (B \xrightarrow{w} Bd)) [r: A \xrightarrow{w} Bd]$
[etc.]

The nested intentionality structure of a macroagent with absolute authority is similar, but differs in one important respect.

- A
- (1) $(A \xrightarrow{w} Bd) [r: A \xrightarrow{w} Bd]$
 - (2) $A \xrightarrow{w} (B \xrightarrow{w} Bd) [r: A \xrightarrow{w} Bd]$
 - (3) $A \xrightarrow{w} (B \xrightarrow{w} (A \xrightarrow{w} Bd)) [r: A \xrightarrow{w} Bd]$
[etc.]

B

- (1) $(B \xrightarrow{w} Bd) [r.A \xrightarrow{w} Bd]$
- (2) $B \xrightarrow{w} (A \xrightarrow{w} Bd) [r.A \xrightarrow{w} Bd]$
- (3) $B \xrightarrow{w} (A \xrightarrow{w} (B \xrightarrow{w} Bd)) [r.A \xrightarrow{w} Bd]$
[etc.]

The first schema brings out very clearly the circularity in the general will. *A* wants something *because B* wants it, but *B* wants it *because A* wants it. Such a circularity does not exist in the latter structure. The difference between the structures is that in the former case we have both

$$[r.A \xrightarrow{w} Bd]$$

and

$$[r.B \xrightarrow{w} Bd]$$

but in the latter case we have only

$$[r.A \xrightarrow{w} Bd].$$

This means that in the authoritarian structure the only kind of reason that exists is a reference to the want of *A*, to Leviathan himself.

The abstract schemas above do not say anything about the concrete social relations which will normally clothe the intentionality structures described. An absolutely authoritarian macroagent, as well as a macroagent with general will, can be based on tradition, i.e. *A* and *B* may be socialized by tradition and come to adopt their respective intentionality patterns in this way. But, in the former case, it may simply happen that a very strong person creates an intentionality pattern within a group by means of his charisma. One can also *imagine* the structure arising via a contract of the type described by Hobbes. To avoid a war of all against all, everyone decides to give *A* absolute power, and make a contract to this effect.

Yet another thing to note is that *both* of the intentional structures I have presented allow that the 'power', of the absolute sovereign or the general will, may be employed in an egoistic or in an altruistic fashion. There is nothing in the schema for the absolute

authority which says *what* he is to will. He can thus satisfy the structure by taking into account either his own welfare or that of everyone else, when he decides what everyone is going to do. The same holds for the general will. The schema above sets out the *form* of the general will. This makes it possible to see that this form is neutral with respect to any questions about the people in whose interest the decisions are made. As I see it, in the discussions of Rousseau's general will, too much weight is usually put on the fact that the general will must make decisions which are for the good of the entire group concerned. The form for the general will can in fact be satisfied in a group which is completely oriented toward *serving one particular person*.

The formal similarity which exists between the two structures discussed is probably of sociological interest. It is a remarkable fact that many religious and political sects whose original ideal was that the group should be a general will and have power-equality among its members, have smoothly turned into strongly authoritarian groups. One aspect of this phenomenon is that the formal similarity between the structures in question helps the members to overlook or repress the real change involved.

I have already pointed out that the structure for the general will is an ideal type but the same goes for the other structures as well. Most actually existing macroagents exist as such only with respect to certain types of action, while I have here described them as existing unconditionally for all types of action. This conditional existence is obvious in the case with organizations and associations, but it is also true of the modern nation-state with its distinction between public and private life. Most macroagents also have *delegated* authority; someone has power via a superior power. I have not taken account of this aspect of macroagents. Moreover, macroagents are often rooted in (the threat of) pure violence. Many microagents become part of a macroagent not because their intentionality fits the intentional structure of the macroagent, but because they are forced to obey the macroagent and be a part of it. One variation of the same theme is obtained when a microagent himself consciously chooses to be part of a macroagent in spite of the fact that his intentional structure diverges from that of the macroagent. He becomes part of the macroagent because he directly benefits from it. His actions are part of the macroagent's, but his soul is not in it.

If we return to ideal-typical macroagents we may note another point of sociological and anthropological interest. Let us first once more write down the *first level* in the 'general-will structure' (1) and in the 'Leviathan-structure' (2), and then in the same notation also the 'will-of-all structure' (3). We get:

	A	B
(1)	$(A \xrightarrow{w} Bd) [r.B \xrightarrow{w} Bd]$	$(B \xrightarrow{w} Bd) [r.A \xrightarrow{w} Bd]$
(2)	$(A \xrightarrow{w} Bd) [r.A \xrightarrow{w} Bd]$	$(B \xrightarrow{w} Bd) [r.A \xrightarrow{w} Bd]$
(3)	$(A \xrightarrow{w} Bd) [r.A \xrightarrow{w} Bd]$	$(B \xrightarrow{w} Bd) [r.B \xrightarrow{w} Bd]$

In the last macroagent it is impossible for the microagents to see themselves only as part of a macroagent. They must *also* see themselves *as microagents*, i.e. they must *as individuals* have content-determined intentions and expressions of will. In the formal structure this is revealed by the fact that A's and B's reasons for their wants are their own will. The first two structures behave in this respect differently. The circularity in the intentional structure of the general will clearly has the contrary character; it hides the fact that a microagent can have a determined will *as a microagent*. In a macroagent with absolute authority, the sovereign must see himself *as a microagent*, but his subjects need not see themselves as microagents; they can see themselves as parts with no will of their own. (The ruler can, of course, see himself as the representative of God, or some such thing, but then it becomes instead a fictive person, God, who must conceive of himself as a microagent.)

This last point is intended to show that *certain* macroagents can in one sense be primary in relation to their constitutive microagents, namely in the sense that the macroagent's specific wants have to be fixed before the microagents' *specific* wants are determined. The microagents are however always ontologically primary, for they constitute the subject-poles in the nested intentionality which constitutes a macroagent. The distinction made between these priority orderings is sufficient for my ontological system to harmonize with the view that, historically, man first saw himself as part of 'non-individualistic macroagents', and only gradually began to see himself as part of 'individualistic macroagents'. All of anthropology testifies to the view that man saw himself as part of a group before he began to see himself as an

individual in the modern sense. In spite of this, the tendency still exists amongst many philosophers to regard the microagent who sees himself as a *microagent* as in *all respects* the fundamental agent. In game theory, to take one example, such a view expresses itself in the belief that 'co-operative games' are always parts of larger 'non-co-operative games'.³³ It is high time for game theory to accept anthropology's results.

15.9 NATURE, MAN, AND SOCIETY

It is now time to summarize the theses of the present theory of categories. In the remaining sections I shall then discuss some epistemological problems. The point of departure for this ontological system has been the difficulty of uniting in a reasonable way the modern conception of nature (nature = 'dead' matter) with the belief that certain concepts which do not refer to such a nature, (e.g. 'action' and 'intention') also refer to something which really exists, and as such demand a place in science.

The conception of nature contained in this system is founded on the fundamental category of space. Space-time is the only entity which fulfils the classical definition of 'primary existents': it exists in and through itself. Everything else is dependent on space-time for its existence. Things cannot exist if they do not exist in container space. The same holds for man and society. Nature, man, and society exist *in* space; and so nature cannot be delimited with the help of the category of space. Any such delimitation has to be made with the help of the intentionality category: *Nature is identical with the non-intentional part of reality.*

This theory of categories diverges from the conception of nature which in modern physics is assumed to imply on two points. First, nature does not consist only of things, substances, (genuine) properties, and relations, but also of tendencies. Second, nature exemplifies universals which are necessarily extended in time. Both organisms and machines are thus, from this perspective, parts of nature. They exemplify the same kind of universals in spite of the fact that their way of coming into existence is completely different. Machines, as distinct from organisms, are constructed by people, they are artefacts; but that is the only ontological difference so long as one assumes that organisms lack intentionality. This equality between machines and organisms is completely clear on the

practical level within modern medicine; I am thinking of the use of artefact organs.

The conception of man presented here takes its departure in the claim that what we call actions are a truly special type of phenomenon. First, they are specific types of four-dimensional universals. In this sense animals, too, can perform actions. This aspect of actions does not tell us what is distinctive about man, but it is none the less important for an understanding of man.

The most important category for the understanding of man is intentionality. This I maintain in spite of the fact that intentionality cannot be identified either with consciousness or with rationality. It is a deeply rooted prejudice that what is most important for the conception of man must be what is characteristic *only* of man. Our modern understanding of man *as a natural* being owes much to the categories employed within physics, chemistry, and technology. The way to a deeper understanding of man *as man* will, I believe, depend on an exploration of the category of intentionality. Whether this will, at the same time, lead to a deeper understanding of the higher animals is in this context a subordinate question; much speaks for these animals having intentionality at least in the form of perception. We understand, literally, mortally important phenomena by having knowledge of our partial identity with dead nature. An increased understanding of the type of intentional structures and patterns I have described may, I hope, in a similar way create preconditions for a better human life whether or not the same structures and patterns exist among other animals.

One specific aspect of intentionality is its ability to be a 'connection at a distance'. That there is such a phenomenon as connection at a distance is actually one of the things which are most obvious in everyday life. But from a modern natural-scientific point of view this appears to be quite mysterious, or indeed occult. As I have drawn the lines of demarcation, this type of connection does not belong to that *part* of reality which nature constitutes. But reality includes intentional phenomena as well. That which is real exists in space and time, and intentional phenomena exist in space and time just as much as non-intentional phenomena. The soul (= the subject-pole in intentional phenomena) occupies the same place in space as does the body. Irreductive materialism solves in an elegant way the mind-body problem. The solution is based on the category of one-sided existential dependence.

I have now mentioned three ways in which the concept of man represents something which cannot be captured in the framework of traditional natural science: (1) actions exemplify universals which are necessarily extended in time, (2) the intentionality category, and (3) the level ontology, which makes it possible for the 'soul' to be placed in space and time.³⁴ Yet another point of the same degree of importance ought to be mentioned. The present ontological system also contains (4) a category of spontaneity. This makes it possible to ascribe to man a certain amount of so-called free will and an ability to transcend what actually exists. Man need not be seen as completely determined by his biological nature and social situation. That man is *partly* so determined, I regard as an undeniable fact.³⁵

It is important to guard against an easy misunderstanding regarding (1) and (2). It is tempting to see *all* real actions as the outflow of a preceding intention ('prior intention'). Such a perspective tacitly presupposes that intentional phenomena always constitute a kind of state, and that they in themselves cannot be actions. However, I have nowhere adopted this perspective. Quite the contrary – it seems to me to be obvious that certain intentional phenomena are actions and in every respect irreducibly extended in time. Many actions such as thinking and deliberating are actions which cannot possibly be regarded as outflows of a 'prior intention'. To try to explain what one is thinking by making a reference to a prior intention to think what one is thinking, leads nowhere. There must exist thinkings which do not have their basis in a prior intention.³⁶ Certainly, one often first has to have an intention in order then to be able to act, but sometimes one must act without intention in order for new patterns to become visible, patterns which when they have become visible can make up the content of intentions.

The intentionality category is not only the key to an understanding of man, but also the key to the understanding of society. Man and society cannot be understood independently of one another. As Marx points out in the famous passage from the sixth of his *Theses on Feuerbach*: 'the human essence is no abstraction inherent in each single individual. In its reality it is the ensemble of the social relations.'³⁷

Can this aphorism be given any reasonable meaning? Can the essence of man exist outside the separate individuals? The only

category which can give meaning to Marx's thought is the intentionality category. The essence of human beings clearly must in some way be anchored in individual people, otherwise it would be meaningless to call it the essence of man. Intentionality and the accompanying distinction between the subject-pole and the intentional correlate can help us here. The *capacity to have intentionality*, i.e. the capacity to be the subject-pole in an intentional phenomenon, must be 'inherent in each single individual', but the human essence perhaps includes some states of affairs which belong to the intentional correlates of presentational intentionality. In that case the human essence does not exist only *in* the single individual.

In order to get a grip on this problem one has to distinguish between two human natures, a natural nature (or biological nature) and a social nature. The former exists by definition only in the spatio-temporal volume occupied by an individual body during its existence. Man's biological nature includes not only such things as genetic structure and the species-specific ability to reproduce, but also the *capacity* to perform or have intentional acts, which of course is not identical with intentionality as such. The latter belongs to an upper ontological level, while the capacity belongs to the substratum.

Presentational intentionality can have both nature and other intentional phenomena as intentional correlates. It is only in the latter case that there are states of affairs containing human properties which partly lie outside a single person. It is only then that the type of nestedness occurs which makes certain states (e.g. shame and pride) wholes which must include at least two participants. This characteristic is most clearly seen in a group with a general will: *A* wants what *B* wants, and *A* wants *B* to want what *A* wants, etc., and vice versa (see the analysis in 15.8). If a determinate content of the will has appeared here, it will be reproduced. The will then exists in both *A* and *B*, but it is nevertheless no will 'inherent in each single individual' but a will which is 'the ensemble of the social relations'. The will can only be understood as a nested intentionality pattern which contains several participants.

This line of thought can also be illuminated by looking at the conventionality of linguistic signs. Today, when a large part of the world's population knows at least a little of a foreign language, the

thesis of the conventionality of the linguistic sign appears to be a truism. Most people understand that the same concept or thought can be related to completely different graphic symbols or completely different sounds. But recognition of this truism can easily obscure an important problem. For a bilingual person the linguistic signs are genuinely conventional. He easily exchanges one type of sign for another. But for the person who is not bilingual, linguistic signs are conventional only in a theoretical sense. They are substitutable in principle, but that person cannot actually make the substitution.

This fact requires an explanation. Why can we not easily exchange linguistic signs in the same way we can change clothes? To a certain extent this non-substitutability can of course be explained in the traditional way by appealing to diverse conditioning mechanisms, i.e. by giving an explanation on the *stratum level*. But may there not also exist an explanation on the intentionality level itself? Let us compare the linguistic situation with a situation involving a general will.

It is not clear from the structure of the general will *what* it is that the general will wills. The will's content is in this sense *conventional*. It is not, however, genuinely conventional for the microagents who normally constitute a general will, but one can imagine a second-order general will – one where those involved not only *have* but also consciously *want* the *whole* of the general will's structure. They see the structure from the outside in the way in which it has been described in the present chapter. If they are able to do this, then they are like a bilingual person. They can choose a certain determined content for the general will, or consciously go from one group with a certain general will to another group with another particular general will. The conventionality in the general will is genuinely conventional and not merely theoretical.

The conventionality of linguistic signs is similar to the conventionality of the content of the general will. There was a time when people quite simply could not imagine languages with signs which differed from their own. Such people, I think, were in the same type of situation as members of a group with a general will where the structure of the general will is not recognized. That this is the case is due to the fact that the linguistic signs are in a nested intentional structure similar to the one the general will has. If a certain graphical pattern is to be a *sign for A*, then it must also be a

sign for *B*, *C*, *D*, etc. And it becomes a sign for *A* because *A* sees that it is a sign for *B*, *C*, *D*, . . . , and sees that *B*, *C*, *D*, . . . see that it is a sign for *A*.

This explanation is very sketchy, but it is an explanation on the intentionality level. I shall not even try to enter into a discussion as to how this sketch is to be filled out in more detail. I want only to explain the way in which it is reasonable to state that a certain language is not 'inherent in each single individual' but is 'the ensemble of the social relations'. A language is a human 'property' which at one and the same time exists both within and without each single person. When one understands the conventionality of the linguistic sign via structures on the intentionality level itself, one obtains at the same time a certain insight into the way a person's nature can be a 'sum' of social relations, where relations are something completely different from external, internal, and grounded relations. In *biological* nature there is a general speech capacity *in* the individual person; but this capacity can only find expression via conventional signs. A *social* nature thereby comes into existence, a language which cannot be said to exist only *in* individual persons.

In principle one can regard the basic human drives, e.g. hunger and sexuality, in the same way as we regarded language. In biological nature there exists not only an *undetermined* linguistic ability, but also an *undetermined* drive for food and similarly an *undetermined* sexual drive. The objects of the drives are, from a biological point of view, *conventional*. This is not a fact which is stranger than the fact that a machine can be lubricated by many types of oil, and nevertheless function perfectly. But this indeterminacy can be determined, partly on the basis of the nestedness of certain intentional patterns. And this occurs in such a way that many biologically possible ways of satisfying the drives become more or less impossible. It is only in very extreme crisis situations that the pattern is broken. Who in our culture can eat raw mice? How many can be sheer will-power change the character of their sexuality? Such questions can be multiplied.

'Above' our biologically given nature there is another, a social nature, and that nature cannot be understood without the concept of '*nested intentionality*'.

I should like to rewrite Marx's aphorism about the social nature of the human essence in the following way:

It is of the human essence to have two essences, one which is inherent in each single individual, and one which is not. The latter, in its reality, is an ensemble of relations of nested intentionality.

The point I have made can be made more simply if one speaks not of *the* human essence, but of *one* person's nature, thereby indicating that different people can have different natures. Biological nature can then be identified with all of the properties and functions of a specific body, while social nature becomes something like the cultural, national, or class identity of the person in question.

Discussions about the relation between the individual and the society are nowadays, more and more, carried on in terms of the three models below (Figure 15.2), which I shall briefly comment on. Following Roy Bhaskar, I shall simply call them Models I, II and III, but for those knowledgeable about sociology, they may be associated in turn with Max Weber, Emile Durkheim, and Peter Berger.³⁸

In the first model, arrows represent intended actions, which does not mean that the intentions are always realized. But if something is to be a social object, it must be the result of meaning-laden human behaviour. What is wrong with the first model is that the link between the individual who acts and the result of his action is too weak. The relation between the action and the result is like that which obtains when someone waters a lawn. The result, the wet lawn, might just as well have been achieved by nature herself. That there was no rain and that therefore some human intention was required to make the lawn wet is an *external contingent aspect of the*

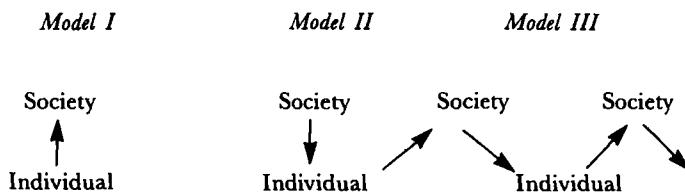


Figure 15.2

result. Social objects and social individuals *are* spatially distinct; and a philosophy without the intentionality category will never be able to connect individuals and their products so as to obtain a social unit. Model I has, moreover, the problem of making social objects too voluntary. There seem to be no societal impediments to action.

The arrow in Model II does not represent the same thing as in Model I. Now the arrow represents a causal relation. Society causes the individual to behave in different ways. Here, too, the bond between society and individual is too weak. Causality is conceived of as efficient causality, i.e. the cause is completely distinct from its effect. In Durkheim himself this is very clear. When he argues for the existence of social facts distinct from psychological facts, he emphasizes strongly that social facts are external in relation to the individual.³⁹ Linguistic signs, currency, and so on – all such things exist outside of me in space. In this, Durkheim is of course right, but he seems to be committing a simple ‘fallacy of composition’. It is possible for each individual in a society to see society as something which exists both outside and independently of him, but it is not possible for all individuals collectively to do this.⁴⁰ To imagine a society completely without people is impossible.

Both Models I and II are of course based on and get their plausibility from a number of true observations. We are sometimes aware that we freely create products which become social products, and we often experience just as clearly the kinds of social constraints which Model II takes as its point of departure. But both models seem to be based on the (mistaken!) classical conception on which intentional phenomena are conceived of as thoughts in the heads of people, and where all relations amongst people, and between people and nature, are either purely spatial or variants of efficient causality. In such a conception one can either emphasize free creation of intentions ‘in the head’, intentions which then, via efficient causality, realize themselves in material products distinct from the intention; or one can emphasize the pressure exerted on the individual person by that which is spatially external. If the society employs pressure, the pressure must be spatially external.

To combine the arrow in Models I and II in order to get Model III solves nothing. I agree completely with Bhaskar’s comment: ‘In seeking to avoid the errors of both stereotypes, Model III succeeds

ONTOLOGICAL INVESTIGATIONS

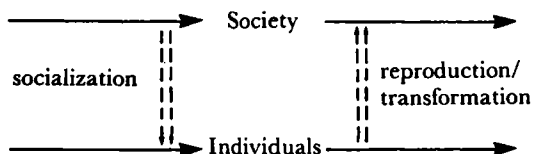


Figure 15.3

only in combining them.⁴¹ Peter Berger, who is one representative of Model III, describes the arrow from the individual to the society as symbolizing 'objectivation', while the arrow from the society to the individual is to symbolize 'socialization'. The problem is that the socialization must take place completely within the single individual, while the objectified product must be placed completely outside the single individual. It once again becomes impossible to understand what constitutes a societal *unit*.

In Berger's case, one finds only external and grounded relations among the individuals, objects and states of affairs which constitute a society. The difference between machines and society turns out to be that society contains a type of parts which, in the space-time they occupy, also contain thoughts.

Bhaskar, who has made critical comments on all three of these models, has himself suggested a fourth alternative which he calls 'The transformational model of the society/person connection'. Graphically, it looks like Figure 15.3.⁴²

This model is, in some respects, an advance over its competitors; but some of the problems I have pointed out for the other models are problems for this one, too. I shall comment first on its strong points and then on its weakness.

According to Bhaskar, it is a mistake to say that man *creates* society; man only 'reproduces' or 'transforms' society. This line of thought can be expressed by saying that for man there always already exists a *societal* matter in the Aristotelian sense. Just as people cannot create statues from nothing, but require a matter to transform, so they cannot create a specific societal institution from nothing, but require some societal matter to *transform*. As a description of how changes actually occur, Bhaskar's perspective is useful. Changes of language, of food-eating customs, of sexual patterns, even large political changes, are better understood as transformations of pre-existing social structures than as totally new

creations. But Bhaskar avoids going into the question about what this societal matter consists of. Such a matter must of course be in some way distinct from natural matter, and it cannot, like nature's prime matter, have existed always. It cannot have arisen historically before man.

Another strong point of Model IV – one which, however, cannot be read off from the figure – is the assertion that, when a social agent produces something, he often reproduces at the same time a certain societal state of affairs. Bhaskar calls this the *dual character of praxis*. Production and reproduction are moments of the same action. When you buy something you produce private satisfaction – if you are lucky – but at the same time you participate in the reproduction of the currency used. When you talk to someone you produce messages and help to reproduce language. If you talk in a new but yet understandable way, you still produce messages but now you help to transform language. Currency and language set limits on what can be done *with their help*, but they do not determine exactly *what* will be done. They can thus be reproduced or transformed at the same time that something specific is produced.

The dual character of praxis has its counterpart in a *duality of structure*. Currency and language, to continue with the same examples, are at one and the same time both condition (or medium) and outcome of actions. Without currency some kinds of transactions would not be possible, but without actions which make use of the currency the currency would not exist. Without a specific language some kinds of communication would not be possible, but without speech acts making use of the language the language would not exist. This kind of conditioning, it is important to note, is both *enabling* and *constraining*. A certain language enables types of communication which would be impossible without language, at the same time as it may make some other types of communication totally impossible. In this respect it is similar to inanimate things. A roof enables you to get shelter from rain but limits the possibilities of sun-bathing at home or of throwing balls up in the air.

Actions (praxis) are performed by persons (agents) and persons have intentions and make deliberations. A person could, most of the time, have acted otherwise than he acted. The same thing, however, is not true of social structures or societal matter in

general. They do not act; they are conditions and outcomes of actions. There exists, Bhaskar says correctly, an 'ontological hiatus between society and people'.⁴³ Bhaskar describes a difference between persons and society, but, unfortunately, he does not probe deeply enough into the connection between agents and societal matter. A specific form of natural matter is very often, like societal matter, a condition for a certain action. But Bhaskar says nothing about the way in which societal matter is distinct from natural matter. From an intuitive point of view, societal matter ought to have a more intimate relationship to agents than does the latter. In spite of Bhaskar's effort, his analysis also means that in the end he is committed to individuals and society standing *against* one another, not *in* one another. (This is also clear from the diagram of Model IV.)

What is the alternative to Models I–IV? I can only repeat what I have now said many times. There is only one way to unite individuals and society, and that is with the help of the intentionality category and different types of nested intentionality. Without nested intentionality there would be no societal matter in contradistinction to natural matter.

Both Berger and Bhaskar are (like myself) partly inspired by marxism. Such an inspiration, however, has its dangers, and this might be the explanation why neither of them gives intentionality the same importance as I do. Marxism is materialistic, and a solution of the problems of the sort I have discussed in terms of intentionality can easily appear to be far too idealistic for a marxist. But such a marxist has not grasped the content of irreductive materialism – and many marxists have unfortunately not done so in spite of the fact that just such an ontology belongs to this tradition. A subject is something with both a *body* and intentionality. And my descriptions of the intentionality patterns connected with the objects of drives and linguistic signs have, I hope, made it clear that both the objects of drives and linguistic signs require material substrata. The material and the intentional are woven together. To satisfy my hunger I need natural matter, but a natural matter which is mediated by my societal nature.⁴⁴ The category of intentionality is fundamental to the existence of man and society, and it does not imply idealism.