

Towards a new approach to the mind-body problem

Master's thesis

Maaïke Harbers

S1217763

April 2006

Advisors:

Dr. F.A. Keijzer

Prof. dr. T.A.F. Kuipers

Prof. dr. M.R.M. ter Hark

**Faculteit der Wijsbegeerte
Rijksuniversiteit Groningen**

Contents

Preface	3
1 Introduction	4
1.1 The mind-body problem	4
1.1.1 Dualism	4
1.1.2 Identity theory	5
1.1.3 Functionalism	5
1.2 Levels of organization	7
1.2.1 Parts and wholes	7
1.2.2 Examples of micro and macro processes	8
1.2.3 A system and its environment	9
1.2.4 Body and mind as a micro-macro relation	11
1.3 Reduction and emergence	11
1.3.1 Reductionism	11
1.3.2 Emergentism	13
1.4 Preview	14
2 Reductionism	16
2.1 Reduction of laws and theories	16
2.1.1 Nagel's bridge laws	17
2.1.2 Hooker and Bickle	17
2.1.3 Kuipers' model of reduction	20
2.2 Reduction of concepts	21
2.2.1 Ontological identity relations	22
2.2.2 Ontological versus epistemological reduction	23
2.3 Micro-reduction	24
3 Emergentism	26

3.1	Three kinds of emergence	26
3.1.1	Nominal emergence	26
3.1.2	Strong emergence	27
3.1.3	Weak emergence	28
3.2	What makes a property emergent?	33
3.2.1	Prediction by simulation	34
3.2.2	Complexity	35
3.2.3	Conclusions	36
3.3	What makes a property non-emergent?	37
	3.3.1 Forms of aggregativity	37
4	Case study	40
4.1	A description of the simulation	40
4.1.1	Cellular automata	40
4.1.2	The simulation	41
4.1.3	Micro and macro levels in the simulation	43
4.2	Reductionistic approach	45
4.2.1	A reduction of the suspicion condition	45
4.2.2	Concept reduction versus theory reduction	46
4.2.3	Conclusions	47
4.3	Emergentistic approach	48
4.3.1	Simulation, complexity and aggregativity	48
4.3.2	Boundary conditions	50
4.3.3	Conclusions	51
5	Conclusion	52
5.1	Reduction, emergence and the simulation	52
	5.1.1 Similarities between reduction and emergence	53
	5.1.2 Differences between reduction and emergence	55

5.1.3 Conclusions

56

5.2 Dualism, identity theory and functionalism 56

5.3 Main question 58

Bibliography 60

Preface

In one of my first artificial intelligence courses was questioned: what actually is intelligence? This was not as easy to answer as it seemed to be. A three year old child who is able to calculate the outcome of eight multiplied by twelve would be denoted as being quite intelligent. Me doing the same thing would not be that impressive. Maybe even no intelligence at all would be described to a calculator performing the same task. What made this difference between human beings and machines? The fact that, in contrast to machines, humans are consciousness and really can ‘understand’ a calculation? Because humans have a mind and machines do not? What actually is ‘a mind’ and why do humans have them? How does ‘the way we are built’ bear on these questions? How do minds relate to their corresponding bodies?

I became very interested in these kinds of problems and started to study philosophy to learn more about them. It seemed strange that we know so little about things we have to do with so much and often deal with so easily. I really enjoyed studying different theories of mind, hearing different opinions on the mind-body problem and having interesting discussions about the topic. However, spending more time on these questions did not provide the answers, it often seemed to make the problems even more complicated. So to conclude my study I did not write a thesis with a solution for the mind-body problem, but this thesis does show some of the results of me working on the problem for some time.

I want to thank Dr. F. Keijzer and Prof. dr. T. Kuipers for being my first and second adviser, respectively. Thanks for all the critical comments, helpful advice, nice cooperation and other contributions to this thesis. I would like to thank Prof.dr M. ter Hark for referee of my thesis.

Maike Harbers

Groningen, April 2006

1 Introduction

1.1 The mind-body problem

One of the most puzzling problems in the philosophy of mind is the mind-body problem, also called the mind-brain problem. This problem concerns the relation between mental and physical properties. On the one hand, the mental seems to be very different from the physical, but, on the other hand, the two also seem to be closely related to each other. The difference appears for example in trying to explain how a conscious experience of pain relates to some firing neurons. It is even more difficult to exactly understand what a strong feeling of happiness has to do with a particular organization of particles. Nevertheless, our thoughts do correlate with events in our brains. We think, feel and experience in some way or other with our bodies or brains. Moreover, our mental activities have physical effects. We have intentions and make plans that at least seem to result in physical actions. So it seems clear that the mind and the body interact with each other in two directions, the question is how they are related to each other in these processes. There are currently three important ways to think about the relation between the body and the mind.

1.1.1 Dualism

In the literature about the mind-body problem, one of the oldest positions is Descartes' substance dualism (Kim 1996). Cartesian dualists think that the body and the mind are two different things made of different substances. The body consists of spatially extended, material substance with volume and weight. Material things cannot be conscious and the material part of human beings (their body) is mortal. The mind is made of non-spatial mental substance: it is invisible, does not occupy volume, weighs nothing and is immortal. The essence of mental substance is thinking or being conscious. The material and the mental both behave according to their own laws and rules, but at the same time they do causally influence each other. In perception for example, the physical stimulation of our sensory system causes us to perceive objects and events around us, in voluntary action our desires and beliefs cause us to move in particular ways.

Although the idea of an immortal 'soul' might seem attractive, there are big problems with substance dualism as an explanation of the relation between the mind and the body. In the first place, why should there be immaterial things in this world? Secondly, if such immaterial things do exist, then how do they interact with material things? In other words, how could

something non-physical, causally influence material bodies that are governed by the laws of physics? Because of these difficulties and the lack of extra explanatory power over other mind-body theories, substance dualism turned out not to be strong enough to survive as a serious theory. Nowadays only very few scientists in the cognitive sciences, psychology and philosophy of mind believe in dualism, in these sciences it is generally assumed that only one kind of substance exists, material substance. So in these present materialistic or physicalistic theories not only our body, but also our mind is a material phenomenon.

1.1.2 Identity theory

An alternative view of the relation between the body and the mind without the problems of dualism is the identity theory of mind (Kim 1996). This position asserts that states and processes of the mind are identical with states and processes of the brain, mental events *are* brain processes. The identity theory is a physicalist theory, which assumes the absence of mental events over and above the neural processes in the brain. Just like ‘lightning’ and ‘electrical discharge’ refer to the same phenomenon, ‘sadness’ and ‘particular neural processes’ also point to the same type of event.

One of the objections against the identity theory is the argument of multiple realizability, an example about pain will help to explain this objection. In the philosophical literature, the neural processes that correspond with pain are traditionally called C-fiber activations. Nowadays the C-fiber theory is an old-fashioned theory, but it still helps to explain the idea of multiple realizability. According to the identity theory, pain *is* C-fiber activation. Consequently, organisms without C-fibers cannot have pain. But it seems implausible that every brain organized in a different way than a human brain cannot have pain by definition. An identity theorist could reply that a more abstract and general physiological description of the brain state of pain is needed, but with this generalization the identity relation is lost. When pain *is* C-fiber activation, it is identical to that particular physical state. But indefinitely many physical states realize pain in all kinds of organisms or systems. So the argument against the identity theory of the mind is that mental states and processes are multiple realizable, there are a lot of possible ways to realize for example pain. For many people this was a great obstacle to the identity theory and a new conception of the relation between mind and brain came up.

1.1.3 Functionalism

The third important approach of describing the relation between the mind and the body is functionalism (Kim 1996). A central point in this approach is the idea of multiple realizable

mental processes and states. Whereas the identity theory says that every mental state has exactly one physical realization (in fact *is* that physical realization), functionalists think that one mental state can be realized in many different ways. According to functionalists, pain is not identical with C-fiber activation, but pain is something with a certain causal relation. Burning a finger for example, hurts and makes one withdraw one's hand. These causal relations of a mental state can be described as a function. The function of pain is preventing organisms from more tissue damage and therefore it plays a role in their survival. What makes something a mental kind is not the underlying structure, but its causal relations and thereby its function. According to functionalists, all mental kinds are causal-functional kinds. All instances of a given mental kind have a particular causal role in common. In this theory mental states are not restricted to humans or animals, according to functionalists everything with the right organization has mental states.

A major problem for functionalism comes from qualia. Functionalists claim that a sensation or experience is *that* particular sensation or experience because of its causal role, but they do not say anything about the experience or sensation itself. Pain might be a tissue damage detector, but pain also hurts. What are the causal relations and the function of the 'hurt'? The functionalist approach cannot give a good account of these qualia, it does not give an explanation of the qualitative features of sensations and experiences.

Another problem with functionalism has to do with the causal power of mental properties. Functionalists define mental properties as properties with a certain causal role. But they also hold that all mental properties are physically realized. The question then is how can mental properties have extra causal powers their physical realizations do not have? Every mental event is a physical event and physical events always fall under the rules and laws of physics. It seems impossible to find a place for the causal role of mental properties, more than just the physical processes in the brain.

The three positions discussed above, all have their own shortcomings and none of them seems to give the right solution to the mind-body problem. With this introduction was aimed to at least make clear the complicatedness of the mind-body problem. The approach to the mind-body problem in this thesis will be different from these three theories: I will investigate the consequences of the mind-body relation viewed as a relation between different levels of organization.

1.2 Levels of organization

1.2.1 Parts and wholes

When different parts participate in a process, they sometimes yield a new possible way of describing the process. Because of their interactions and the way they are organized, the parts together form a whole with its own behavior. Besides a description on the level of the parts (the lower level), a description on the level of the whole formed by the parts (the higher level) originates. The latter often involves properties or phenomena, which cannot be found on the lower level. Examples of phenomena with different levels of organization are fluidity, a swarm of birds, a tornado and life. All of these things as a whole possess properties that the parts the whole consists of do not possess. But at the same time, the parts are necessary to form the whole. Water is fluid, but one single water molecule is not. Every bird in a swarm just flies a certain route, but all the birds together show a more complex flocking behavior.

In systems that are organized in different levels, it often makes sense to talk in terms of micro and macro processes and objects. Micro processes are the processes of the micro objects, the parts that constitute the ‘lower level’. Macro processes are the processes of the whole system on a ‘higher level’, which arise out of the interactions of the parts. In this thesis, the terms micro and macro will refer to the respective level of organization. However, the terms are relative: a system with micro and macro properties is often part of a larger organization, in which the system is a component of a bigger system. Its interactions with other parts in that ‘higher order’ system can then be viewed as micro processes, the macro level (the whole system built up of parts) becomes a micro level (a part of a bigger system). Parts of a system can also be built up of smaller parts and can be analyzed as whole systems with their own macro properties, here the micro level becomes a macro level.

These micro-macro relations create so-called levels of organization, in which all levels (except a possibly lowest base level) are micro and macro levels at the same time. Wimsatt and Kim both referred to these compositional levels. Wimsatt said the following:

“Hierarchical divisions of stuff, organized by part-whole relations, in which wholes at one level function as parts at the next levels.” (Wimsatt 1997)

Kim called the compositional levels:

“Different layers of the world.” (Kim 2002)

Kim (2002) also mentioned different sciences in the context of levels of organization. According to him, the topics of investigation in different sciences can be seen as different layers of the world, related to each other as micro and macro levels. Physics, chemistry, biology, psychology and sociology all investigate the world around us, but they focus on different aspects of it. The research objects of one science are the parts out of which the topic of the next science exists. The cells investigated in biology for example, are macro level objects in relation to what chemistry is about, but they form the micro level objects in psychological research. So the status of a system or object (micro or macro phenomenon) depends on the specific focus and context of interpretation.

1.2.2 Examples of micro and macro processes

The first of three examples to illustrate the concept of processes or phenomena with different levels of organisation is the organization of an anthill (Pfeifer & Scheier 1999). All the single ants in an anthill execute simple jobs and no ant or group of ants thinks about what has to be done in the anthill as a whole. No one has an overview over the whole group or gives all the orders. But still, all the ants together placed in their environment form a very complex and well-structured organization. There is food for everyone, the hill is kept in repair, the newborn ants are taken care of and enemies are attacked. This is possible because of the interaction of the ants with each other and their environment.

The example of the anthill assumes a certain simplicity of the actions of the ants, no rich intellectual worldview is attributed to them. This is a very plausible assumption, but the exact rules according to which the ants behave are not known. For this reason computer simulations are involved in research of these phenomena. Such computer simulations do not depend on plausible assumptions or on real world observations, the initial conditions and the rules of processing are exactly known with one hundred percent certainty. Another advantage of computer programs is that simplifying problems and alternating the variables in a model is very easy. This makes computer simulation a quick and useful tool for investigating what happens on the macro level, when parameters on the micro level are changed.

The second example of phenomena with a micro and a macro level is about the flocking behavior of birds, which is very successfully simulated in a computer program by Reynolds (1987, cited in Pfeifer & Scheier 1999). The simulation shows exactly the same kind of behavior as real flocks of birds. All the birds in the simulation follow only three rules: stay close to birds nearby, avoid collisions and match velocity with the birds in your neighborhood. These simple

micro level rules produce the flocking behavior of birds on the macro level. There is no macro level rule like ‘stay together in the group’ or ‘follow the leader’ or ‘always fly to the south’, but the flock does show organized behavior. All the birds stay together and fly in the same direction. Although communication might seem necessary to be able to produce such difficult behavior, in computer simulations the three micro level rules are sufficient to create the flocking behavior on the macro level.

The last example is called *The Game of Life*, devised by the mathematician John Conway (Dennett 1991), also a computer simulation. The Game of Life is not a real game, it is ‘played’ on a two-dimensional grid divided into square cells. Each cell is ‘on’ or ‘off’ at a certain moment and the states of the cells change over time. Time in the Game is discrete and the state of a cell depends on the states of its neighbor cells in the previous iteration. Conway’s ‘rule of life’ below, is about the conditions under which the cells change their state.

“Each cell, in order to determine what to do in the next instant, counts how many of its eight neighbors are ‘on’ at the present instant. If the answer is exactly two, the cell stays in its present state in the next instant. If the answer is exactly three, the cell is ‘on’ in the next instant whatever its current state. Under all other conditions the cell is ‘off’.”

(Conway, cited in Dennett 1991)

This micro level rule determines all the action on the grid and starting from an arbitrary begin state, it can create very complex, unexpected patterns. An example of such a macro pattern is a glider, which is a five-pixel configuration that ‘moves’ across the field. Other examples are shooters and eaters, configurations that ‘shoot’ and ‘eat’ other cells that are ‘on’. These patterns cannot be found by just looking at the cells and its rules, an overview of the whole is needed.

In all of these examples the parts behave according to simple rules: the ants, the birds and the cells on the grid. All the examples also involve a higher level process: the whole organization of the anthill, the flocking behavior of the birds and the shooters and gliders in the game. It is claimed that the behaviors of patterns in the higher level cannot be found in the rules at the lower levels, although nothing else arranges the higher level processes.

1.2.3 A system and its environment

Changing systems with parts that interact with each other as described in the previous subsection, often, if not always, influence and are influenced by their environments. A relatively young approach in the cognitive sciences, called *interactionism* (Keijzer 2001), focuses a lot of attention

to the environment. It originated out of criticism on the basic assumptions of traditional cognitive psychologists in their research of behavior. The traditional approach would focus too much on the processes inside a human head and external factors would be neglected. The 'new approach' gives a better account to the circumstances in which a cognitive system operates and it has gained increasing attention over the last fifteen years. Two main notions of interactionism are *embodiment* and *situatedness*. Situatedness refers to the environment in which a cognitive system operates; environment plays an important role in the behavior of a system and the content of its mind. The second aspect, embodiment, refers to the importance of a physical body for intelligent actions. An agent's capabilities for sensing and moving partly determine its possibilities for solving a problem. The importance of a system's body and environment is more and more accepted in the cognitive sciences.

Bechtel (2003) states that to fully understand a system it is necessary to interpret the system on the macro level. According to him, looking to the macro level means to look at the environment of a whole system. If the system itself is a component of another system, it is also necessary to look at that bigger system to understand the behavior of its component. Bechtel says that boundary conditions on the macro level can constrain lower-level processes. For example when someone moves his arm, all the cells in his arm also move because they are part of the arm. It is possible to analyze what happens with a cell only on the micro-level, but Bechtel stresses that to fully understand a mechanism you have to look upward, downward and sideways.

As stressed by Bechtel and by interactionism in general, I also think the environment plays an important role in the way organisms and systems behave. Embodiment and situatedness often appear together, and for organisms and systems with a body interaction with their surroundings is unavoidable. By looking at the environment of a system, a better understanding of the system's macro level is reached. The environment also plays a role in finding macro properties in a system. We recognize and understand the macro properties in a system, because we see the system as a whole within a specific environment. An example is the macro property of a car: a car can move and its parts cannot. The environment plays a crucial role here, without an environment the car would not be able to move at all. Another example is the fluidity of water, we know that water is fluid because we perceive water within an environment also consisting of non-fluid things. So because we perceive systems within an environment, we view them as a whole and attribute particular properties to them. Without an environment, speaking of macro properties would not be possible.

1.2.4 Body and mind as a micro-macro relation

In an interpretation of the relation between the mind and the body as a relation between a micro and a macro level, the parts on the micro level could be formed by neurons in the brain, the objects of investigation in the neurosciences (the body). Because of the interaction between all the different neurons and their environment, new macro processes arise. This new ‘whole’ could be what we normally call thinking, feeling, sensations, consciousness, experiences or just mental activities, the things that are subject of psychological research (the mind). As in the examples above, there is not ‘a someone’ who thinks of what will happen on the macro level in this interpretation. We do not have a homunculus in our brain governing all the brain processes, the micro processes just go their way.

In this thesis I will investigate the consequences of this interpretation involving different levels of organization to the mind-body problem. By viewing the mind-body problem from this micro-macro perspective, nothing but (neuron) cells on the micro level are involved in the mind-body question, matter is the only substance. Macro phenomena are created because of the interactions of the cells and no extra ‘stuff’ has to be added to explain mental events. This position is a physicalist one, and therefore not compatible with dualism. So because of the micro-macro level approach of the relation between the mind and body in this thesis, dualism is dismissed from consideration. The approach leaves the possibilities of the identity theory and functionalism open, and this issue will be placed in a wider context: a discussion between reduction and emergence.

1.3 Reduction and emergence

In the literature, reduction and emergence are the two major ways of looking at the relation between higher and lower order levels. Investigating these general approaches to the relation between micro and macro phenomena, could lead to more insight in the relation between the mind and the body.

1.3.1 Reductionism

Until recently, reductionism was not a very well represented position in the philosophy of mind, most people supported non-reductionistic theories. Two well-known reductionists made the following statements, which indicate a negative general attitude towards reduction. Jaegwon Kim wrote the following about it.

“If you want to be politically correct in philosophical matters, you would not dare come anywhere near reductionism, nor a reductionist.” (Kim 1998, p 89)

John Bickle started his book with this statement.

“Consensus holds that reductionism is dead.” (Bickle 1998, p1)

Despite this negativity, reductionism also has its proponents and it is even regaining popularity the last years. I will ignore Kim’s and Bickle’s statements and treat reductionism as a serious position, just as they did themselves.

An important person in the history of reductionism is Ernest Nagel, he formulated an account of reduction in the 1950s (Nagel 1961), which had a lot of influence in many of the discussions about reduction after that. Nagel’s model of reduction mainly dealt with the reduction between different scientific theories. The aim of his model was to reduce theories that described the same phenomena in different ways. Besides theories, other reductionists made laws, concepts, processes and phenomena also subject of reduction. According to reductionists, if different ways of describing the same thing are possible, a description on a lower, more basic (micro) level is always preferable above a macro description. In their view macro processes do not 'have' or 'do' something, which cannot be found in the micro processes. The claim is that macro phenomena are nothing more than: the parts on the micro level, the rules according to which the parts behave and their interactions. A reduction then, is a derivation of the macro description from a suitable micro description, it is said that the macro phenomena are reduced to micro phenomena (Kuipers, personal communication).

The identity theory in the mind-body debate is closely connected to the notion of reduction. According to identity theorists, the physical and the mental are identical, they think that theories of the mental can totally be reduced to physical theories. Theories about the mental are only used for practical reasons. An even stronger reductive position in the philosophy of mind is called eliminative reductionism. Paul and Patricia Churchland for example, think that when psychological theories are successfully reduced to neuroscientific theories, the psychological theories can totally be eliminated (Bechtel et al 2001). They claim that even for practical reasons the psychological theories are no longer needed then, neuroscientific theories are sufficient to describe the mental phenomena.

1.3.2 Emergentism

The first emergentists stem from the second half of the nineteenth century, their view is now known as *British Emergentism*. This view originated in the context of philosophy of science, the philosophers tried to make distinctions between emergent and resultant laws and effects (Stephan 1992). George Henry Lewes belonged to the British emergentists and he was one of the first philosophers who mentioned emergence. He gave the following description.

"There are two classes of effects markedly distinguishable as resultants and emergents. Thus, although each effect is the resultant of its components, the product of its factors, we cannot always trace the steps of the process, so as to see in the product the mode of operation of each factor. In this latter case, I propose to call the effect an emergent."
(Lewes 1875, cited in Stephan 1992, p27)

Another British emergentist, C.D. Broad, gave the following explication of the concept of emergence. According to him, emergent theories claim that

"the characteristic behavior of the whole could not, even in theory, be deduced from the most complete knowledge of the behavior of its components, taken separately or in other combinations, and of their proportions and arrangements in this whole." (Broad 1925, cited in Beckerman 1992, p16)

Both of these citations emphasize the unpredictability of macro states, even given complete information of the states and the rules on the micro level. Not everybody agrees with this classical concept of emergence. Andy Clark, for example, a philosopher of this time, sees emergence as the process by which complex, cyclic interactions give rise to stable and salient patterns of systemic behavior.

"Emergent phenomena are often precisely those phenomena in which complex interactions yield robust, salient patterns capable of supporting prediction and explanation." (Clark 2001, p117)

This contradiction between the unpredictability or predictability of emergent phenomena already shows some of the controversy about the concept of emergence. Several other points of discussion also lead to disagreement among different emergentists, many different forms of emergence are propagated. The basic idea of emergence, common to all kinds of emergentism, is the emergence of higher level phenomena out of lower level ones. The systems or processes,

which play a role in the higher level, are formed out of parts, which are the objects of the lower level. All emergentists attribute properties to the macro phenomena, which micro level phenomena do not possess. Mark Bedau (1997) distinguished two hallmarks of emergence, which hold for a lot of different notions of emergence.

- 1 *Emergent phenomena are dependent on underlying processes*
 - 2 *Emergent phenomena are autonomous from underlying processes*
- (Bedau, 1997)

These two criteria summarize the common aspects of many different accounts of emergence. Later I will show more extensively what Bedau meant with these criteria and how they structure and unify various notions of emergence.

In the mind-body debate the notion of emergence can be associated generally with functionalism. This connection is less often made and maybe less evident than the connection between the identity theory and reductionism. But although emergentism and functionalism are not the same, they do have important features in common. Functionalists say that the exact realizations of mental phenomena do not matter, as long as a specific function is performed. According to emergentists, macro phenomena have certain properties, which are not present on the micro level. The existence of the macro properties depends on general structures of the micro level, not on specific configurations. So functionalism and emergentism agree on the independence of realization of macro properties. In functionalism these properties are always mental properties, emergence is broader and applicable to more macro phenomena.

1.4 Preview

A clear, unambiguous and generally accepted solution of the mind-body problem in the philosophy of mind is not reached and maybe never will be. Maybe the whole material is too complex for us to oversee and we are not able to understand it. But people not only interpret the relation between the mind and the body in different ways, this also happens in more open and simple systems with different levels of organization, which are much easier to oversee and understand. I think it is strange that, although people have the same data or knowledge at their disposal, they draw such different conclusions from it. Points of discussion are for example: what makes a macro phenomenon a macro phenomenon? Can macro phenomena be reduced to micro phenomena? Are macro processes real? Do they just originate in our interpretations of the world?

Reductionists and emergentists will give different answers to these questions and reductionism and emergentism are often perceived as two very different, maybe even contradicting, theories. But despite a lot of differences, the theories do not totally contradict each other and have some features in common. In this thesis, besides describing the differences, I will also focus on similarities between reductionists and emergentists or identity theorists and functionalists.

The main part of the thesis will focus on the different ways to describe the relation between different levels of organization, to investigate differences in interpretations of emergentists and reductionists. After considering more simple processes with a lower and a higher level, I will return to the mind-body problem in the conclusion and apply the results of my research to the mind-body problem. Taken all this together, the threefold main question of the thesis will be this.

(a) To what extent are higher order processes real and autonomous with respect to their underlying realizations, (b) do reduction and emergence describe the relation between the micro and the macro level in a different way, and (c) how does this bear on the mind-body problem?

To answer these questions, I will discuss reduction and emergence, in chapter two and three respectively, and try to find some concrete conditions and features that characterize the two approaches. The subject of chapter four will be a case study of a simple computer simulation with higher and lower order properties. The theories of chapter two and three will be applied to analyze this simulation. In the final chapter, the conclusion, the differences and similarities between emergence and reduction will be discussed in the context of the case study in chapter four, and in a more general way. Finally I will return to the mind-body problem and discuss my conclusions, among others by comparing my approach with the three theories I described in the first section of this chapter.

2 Reductionism

Reduction is a very broad concept and it has a lot of different forms and applications. But in every reduction 'something' is reduced to 'something else'. In considering different types of reduction, two important aspects can be distinguished. The first aspect to be analyzed is the subject of reduction: what are the 'to be reduced' and the 'reducing'? The second aspect is the reduction itself: what does the reduction look like?

With the first aspect, the subject of reduction, I mean the kind of things that are reduced. Theories and laws are reduced to other theories and laws, but concepts can also be the subject of reduction. The 'to be reduced' can be on the same level of organization as the 'reduced', which is called intra-level reduction, or they can be on different levels, called inter-level reduction.

The second aspect that varies among different reductions is the type of reduction, the way something is reduced. A reduction can be eliminative, in that case the 'to be reduced' is totally eliminated in favor of its reduced equivalent. The opposite of this is non-eliminative reduction, here the 'reducing' and the 'to be reduced' keep on existing together and are viewed as different approaches of the same phenomena. Other kinds of reductions are for example micro-reductions, identity reductions and reductions with bridge laws. The meanings of these concepts will be explained later in this chapter.

Because reduction has so many different forms, splitting up the topic will make it easier to manage. Although I think the distinction between 'subject of reduction' and 'type of reduction' as made here is a legitimate one and will help to better understand reduction, it is impossible to discuss the two aspects separately. The type of reduction is closely interwoven with the subject of reduction. Therefore I chose to split up the discussion of reduction in the reduction of theories and the reduction of concepts. The reduction of laws and theories gives the opportunity to give an overview and discuss several kinds of reduction in a structured way. The reduction of concepts is important because it will mainly be used in the rest of this thesis.

2.1 Reduction of laws and theories

In this section I will first discuss Ernest Nagel's account of reduction, and then continue with critics on his theory and introduce further developed theories of theory reduction of Hooker,

Bickle and Kuipers. In all of these accounts, reduction means the reduction of a theory T1 to another theory T2, unless otherwise stated.

2.1.1 Nagel's bridge laws

Nagel's model of inter-theoretic reduction has dominated the discussion of reduction during the last decades (Marraffa & Nani 2002). According to him, reduction is a kind of explanation and Nagel's focus is on the derivation of laws. In his model, laws of a target theory Tr are derived from those of a base theory Tb. To connect the predicates of the two theories, sometimes cross-theoretic connection principles are needed. These connection principles are called bridge principles or bridge laws. A simple scheme without the details of his account looks like this.

Tb (base theory) & **BP** (bridge principles if necessary)

Logically entails

Tr (target theory)

(Marraffa & Nani 2002)

Nagel called the reductions without bridge laws homogeneous reductions. But he primarily focused on cases where bridge laws are needed, termed heterogeneous reductions (Looijen 1998). A classical case of reduction with partial successes is the reduction of thermodynamics to statistical mechanics. Nagel's model of reduction applies very well to some parts of the theories; with the aid of some simple bridge laws, some thermo-dynamical laws are reduced to mechanical laws (Beckermann 1992, p 107).

Nagel's model of reduction does not apply that well in all cases of reduction. A lot of people criticized his account on several points. But although many people did not agree with all of Nagel's ideas, they explicated why and on what points. A lot of new accounts of reduction were different from, but also based on Nagel's classical model. That is why his account is so important.

2.1.2 Hooker and Bickle

One of the problems with Nagel's account of reduction is the implication that if Tr is wrong, Tb must be wrong too. This is not in agreement with a lot of familiar examples of reduction, which do involve wrong Tr's. Clifford Hooker developed a new scheme of reduction, which deals with this problem.

Tb (base theory) & **BC/LA** (boundary conditions or limiting assumptions if needed)

Logically entails

Ib (image of Tr in terms of Tb: a set of theorems of (restricted) Tb)

Which is relevantly isomorphic to (“analogous to”)

Tr (target theory)

(Marraffa & Nani 2002)

This scheme does not imply a wrong Tb if Tr is wrong. Hooker introduced an Ib (image of Tr), which is completely characterized within the framework and vocabulary of Tb. A step has to be made to go from Ib to Tr, the difficulty of this step varies among different reductions. His approach yields a spectrum of reductions, which goes from smooth to bumpy reductions (Bickle 1998). The amount of corrections that have to be made to Tr, determines the smoothness of a reduction. In smooth reductions Ib and Tr closely fit, in bumpy reductions more corrections have to be made. Hooker’s account differs from Nagel’s account in the possibility of bumpy reductions. With Hooker’s model, Tr can turn out to be wrong and replaced by Tb. Nagel only took the smooth reductions into account, the situations where Tr and Tb both are correct theories.

Figure 1 below shows Hooker’s spectrum of smoothness, smooth reductions are located on the right side and bumpy reductions on the left side of the spectrum (Bickle 1998). Very smooth reductions correspond with identity relations, no corrections in the theories have to be made. Very bumpy reductions eliminate the reductive theories, they are totally replaced by the base theory. The smoothness or bumpiness of a reduction is a continuous aspect, every reduction has a location somewhere on the spectrum. An example is the reduction of visible light to light with a certain wavelength, this reduction is located on the left side of the spectrum. The replacement of phlogiston by other theories of combustion is a bumpy reduction located on the right half of the spectrum. The top line, from perfect smoothness to limit bumpiness, indicates the smoothness of theory reduction. The bottom line refers to the ontological consequences dependent on the inter-theoretic reduction. The relation between the two lines is the step from Ib (the image of the target theory in terms of the base theory) to Tr (the target theory). Ib and Tr have to be relevantly isomorphic (analogous) to each other.

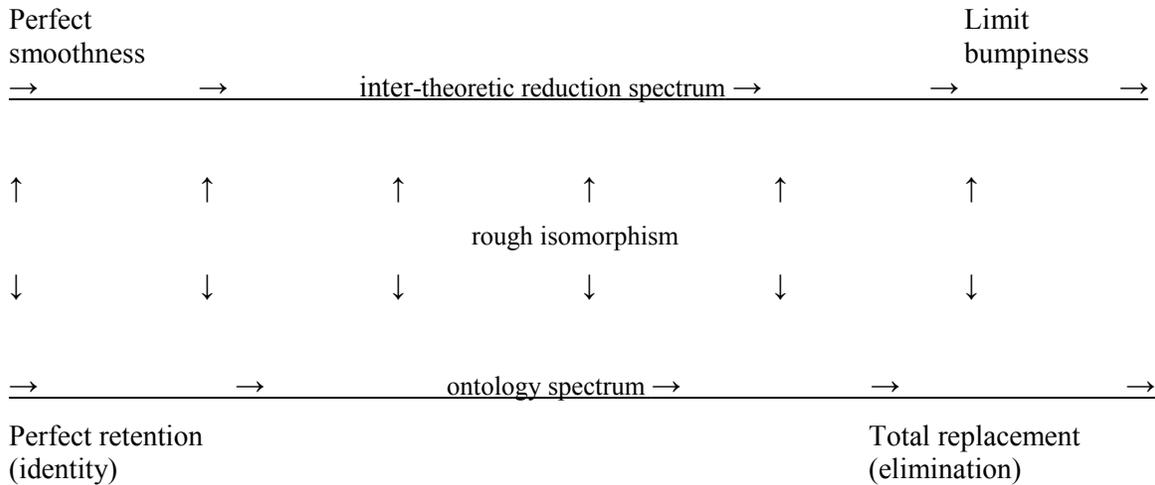


Figure 1: Hooker’s spectrum of reductions (Bickle 1998, p 30)

John Bickle is positive about Hooker’s ideas, but he thinks the theory is incomplete. Hooker is not capable of locating reductions on the spectrum, because there are no tools for measuring their exact smoothness. And because he cannot make the distinction on the theoretical level (the top line in the figure), he neither can draw any ontological conclusion from a reduction (the bottom line). Bickle wants to solve this problem in his own account of reduction: *New Wave Reduction*. According to Bickle, inter-theoretic reduction is the potential constructability of an image of Tr within Tb. His account is based on structural similarities between Tb and Tr. If the structures of both theories resemble, Tr is reducible to Tb. To assure that Tr and Tb, besides enough similarities, also have ontologically something in common, Bickle adds the constraint of an ontological reductive link (ORL). The exact details of Bickle’s theory are very complex, but his theory provides a way to locate reductions on Hooker’s inter-theoretic spectrum of reduction.

Bickle applied his New Wave theory of reduction to the mind-body problem. Because his reduction model is a model for reduction between theories, he reformulated the traditional mind-body problem as an inter-theoretic problem. Bickle thinks the mind-body problem can be solved by reducing psychological theories to neurobiological theories. In the first chapter I formulated the mind-body problem not as an inter-theoretic problem, but as a problem of the relation between micro and macro phenomena. This is a very different approach and Bickle’s account of reduction will not be applicable to my formulation of the mind-body problem. Kuipers’ account of theory reduction involves, besides inter-theoretic reduction as Bickle sees it, also the special case of the inter-theoretic reduction of macro level theories to micro level theories.

2.1.3 Kuipers' model of reduction

Nagel only discerned heterogeneous and homogeneous reductions: reductions with and without bridge laws. Besides these non-eliminative identity reductions, Hooker added the possibility of eliminative, bumpy reductions in his theory. Theo Kuipers (2001) developed a model of inter-theoretic reduction in which he also reserved a special place for micro-reduction, a form of reduction not explicitly mentioned by the preceding philosophers. Before the focus on micro-reduction, I will give a short global account of Kuipers' theory. He discerned the following five possible steps, which could be part of the explanation of laws by theories (Kuipers 2001).

- 1 *Application.* All explanations and hence reductions start with an application step, in reductions this leads to the application of the explanatory (reducing) theory to the system to which the law or theory to be explained is applied. This is the only step present in all reductions, the other steps are possible steps.
- 2 *Aggregation.* This step applies when there is a part-whole relationship between the objects of the two theories. The aggregation step results in an aggregated law, which explains the phenomena of 'the relevant number or set of parts' (the whole) in terms of the level of the parts.
- 3 *Identification.* If this step occurs the reduction is heterogeneous, T1 and T2 have different vocabularies. Terms of T1 are ontologically identified with terms of T2 and such bridge laws or rules of correspondence are required for reduction.
- 4 *Correlation.* As with identification, this step also involves a heterogeneous jump of language. But instead of ontological identities, what is referred to by the terms in T1 and T2 are causal relations.
- 5 *Approximation.* This final possible step is the only non-deductive step, in which some mathematical or logical approximation device simplifies a deductively derived law. The approximation step gives the opportunity to get rid of superfluous terms.

Only the first step occurs in every reduction of laws and theories, the other steps are possible parts of a reduction. To speak of a reduction the first step and at least one of the steps two, three or five have to be taken. Step three and four together are more generally called the transformation step. Different combinations of steps lead to the kinds of reduction mentioned below, a reduction including all five steps is a heterogeneous, approximative, micro-reduction.

Heterogeneous reduction (includes at least step 1 and 3) is reduction with an identification and possibly a correlation step, reduction without these steps is homogenous. This

kind of reduction corresponds with Nagel's heterogeneous reduction. In Nagel's vocabulary the transformation rules are rules of correspondence or bridge principles.

Reduction with an approximation step is *approximative reduction* or corrective reduction (includes at least step 1 and 5). Reduction without such a step is called deductive. This type of reduction corresponds with the non-smooth to bumpy reductions on Hooker's spectrum of reduction.

The last kind of reduction Kuipers discerns is *micro-reduction* (includes at least step 1 and 2), reduction with an aggregation step. Reductions without such a step are termed iso-reductions. Micro-reduction is about two different descriptions or theories of phenomena, happening on the same time and in the same space, one on a macro and another on a micro level. Nagel's model of reduction does not give attention to micro-reductions, neither do Hooker's and Bickle's accounts.

Micro-reduction seems to fit best to my approach to the relation between the mind and the body, it is the only kind of reduction involving part-whole relations. However, micro-reduction here is described within a model of inter-theoretic reduction. And in contrast to Bickle's formulation of the mind-body problem, I want to look at the relation between the mind and the body as a micro-macro relation without insisting that theories have to play a role. In my opinion the necessity of theories in Bickle's approach of the mind-body problem is one of its weak points. Because of my way of formulating the mind-body problem, the next section will pay attention to the reduction of concepts. The reduction of concepts (phenomena, processes, properties) takes place without the participation of theories. After a general approach of concept reduction, in the last section of this chapter I will discuss the *micro-reduction of concepts*, the most proper kind of reduction to apply to the mind-body problem as formulated in this thesis.

2.2 Reduction of concepts

The reduction of concepts is closely related to the reduction of theories and laws, concepts are part of theories. The reduction of laws presupposes the reduction of concepts. The other way around it is different, the reduction of a concept does not imply the reduction of a theory or law. And in contrast to the reduction of laws, the reduction of concepts is often not seen as a form of explanation (Kuipers 2001).

2.2.1 Ontological identity relations

Rick Looijen is strongly influenced by Kuipers' work, which is currently best presented by *Structures in Science*, and he provides a critical extension of Kuipers' position on reductionism. Looijen (1998) gives his own definition of concept reduction.

“Concept reductions are relations (ontological identity relations) between two different concepts or representations of one and the same thing or of one and the same attribute.” (Looijen 1998, p 69)

In Looijen's definition, a proper concept reduction includes an identity relation between two different concepts or representations. In the literature, two kinds of identities are being distinguished (Looijen 1998). The first is a token-token identity involving two identified tokens, the second is type-type identity with two identified types. A type is an abstraction of something, a kind, and a token is an instantiation of a certain type or kind. An individual dog is a token of the type 'dog' and a particular table is a token of the type 'table'. An example of a *token identity* is the planet Venus, which was referred to by the 'morningstar' and the 'eveningstar'. People first thought these were two distinct stars, but later they discovered it to be two appearances of the same planet, Venus. Examples of *type identities* are the identity of 'temperature' with 'mean kinetic energy of molecules' and that of 'pressure' with 'kinetic pressure', that is, 'average of the momenta transferred from the molecules of the gas to the walls of the container'. In concept reduction, reductions are always type identities.

According to Looijen (1998), concepts of a thing and concepts of an attribute can have reduction relations. Thing reductions and attribute reductions are both type identities. *Thing identities* are identities of a thing or object, for example the identity between 'light' and 'electromagnetic waves' (Looijen 1998). An *attribute identity* is a general identity in which properties, relations or functions are identified. The type identity of macroscopic pressure of a gas being the same as the kinetic pressure of the gas molecules given in the previous paragraph is an example of an attribute identity. According to Kuipers (2001) and Looijen (1998) the reduction of attributes runs analogous to the reduction of things.

Another variable of reductions is whether the reduction is singular or multiple (Kuipers 2001). In *singular type reduction*, exactly one type is identical to another type, for example 'temperature' and 'the mean kinetic energy of molecules' (Kuipers 2001). If a particular type is reducible to (has ontological identity relations with the union of) more types, it is a *multiple type reduction*. This applies in the case something is multiple realizable, when several instantiation types of one type to be reduced are present. An example of a multiple realizable kind is vision,

this is accomplished in many ways among different animal species. All the realizations lead to the same property, the ability to see.

2.2.2 Ontological versus epistemological reduction

Ontological identity relations of concepts are often opposed to causal correlations of concepts. Correlations are subject to causal explanation, they are causal relations, which call for a deeper, causal explanation (Looijen 1998). Identities cannot be causally explained, they are ontologically identical because the concepts or representations involved refer to the same ontological thing or attribute. This implies that two identical concepts are on the same ontological level and that the difference between the two solely exists in epistemological terms. Because Looijen claimed that concept reductions must involve ontological identity relations, concept reduction is always *epistemological reduction* according to him (Looijen 1998).

Ontological reduction refers to the reduction of a concept of one ontological level to a concept of another ontological level. According to Looijen this is never possible if an ontological identity relation is required for reduction, two concepts on different ontological levels can never have an ontological identity relation. He calls the expression ‘ontological reduction’ misleading and a contradiction of terms (Looijen 1998). Kuipers has a different opinion, he thinks it is ‘tempting to formulate a plausible explication of the idea of ontological reduction of concepts’ (Kuipers 2001, p150). In the next section I will describe his account of micro-reductions, in which macro level concepts are reduced to concepts in micro level terms. Kuipers (2001) and Looijen (1998) agree on the claim that two ontological levels are involved in part-whole relations, but Kuipers calls the (micro-)reduction of the whole to the parts an ontological reduction, Looijen does not.

2.3 Micro-reduction

In his discussion about the micro-reduction of concepts, Kuipers (2001, p142) introduces the term *aggregate concept*. An aggregate concept is by definition defined in terms of micro-types of micro-objects for aggregates of such micro-objects. An example Kuipers gives of an aggregate concept is ‘mean kinetic energy’, the definition of temperature in terms of properties of its micro-objects. By representing macro concepts this way, the microstructures of macro concepts become clear and the micro and the macro concepts can (approximately) be identified with each other according to Kuipers (2001).

Besides the question whether the micro-reductions of water given above are ontological or not, I think that the reductions are not satisfying and something is missing. The properties, relations and functions of water do not appear in Kuipers' and Causey's formulations. The smallest possible amount of water is not transparent or fluid, but an aggregate of H₂O is. The formulations above do not explain how the aggregate of molecules becomes transparent and fluid, although a single H₂O molecule is not. Just putting the single parts of a car on a pile does not make them a car with the ability to move, something more is necessary. Analogous to this, more is needed in the reduction of water to H₂O molecules than just saying that water is a lot of H₂O molecules together.

In this example of the reduction of water to H₂O, more attention should be paid to the macro properties of water instead of focusing on its components. An amount of water consists of a number of H₂O molecules and nothing more than that, no extra parts are added. A system per se is not more than the sum of its parts. But although the amount of water just consists of a set of H₂O molecules, it has properties, which the H₂O molecules do not have. The properties of water are more than the sum of the properties of its parts. William Wimsatt, who analyzed the different ways a system can be put together, said: "there is no point in speaking of systems *in toto* as being or not being aggregates of their parts, it is better to talk about properties of systems and their parts" (Wimsatt 1986). I agree with him that a better way to reduce water to H₂O, is by reducing the macro properties of water to properties of the H₂O molecules. The reduction of transparency for example should deal with properties of the structures of H₂O molecules, which explain why water is transparent. Because of the specific microstructures and the interaction between H₂O molecules, an aggregate of H₂O molecules yields transparency.

Concluding I think that micro-reduction does not ontologically reduce a macro level to a micro level. Micro-reduction as discussed in this chapter is the epistemological reduction of a macro level description to a micro level description. Furthermore, if aggregate concepts are used for micro-reduction, they should be about macro properties instead of describing by which parts a whole is built up. Micro-reduction should explain the way a macro property of a system arises, depending on the properties of its parts. By explaining the way something is built up, the macro property is explained in terms of micro properties and thereby a reduction is given. The formulation of aggregate concepts is a way to accomplish this, as long as enough attention is paid to the properties of a system instead of systems per se.

In the rest of the thesis I will use the terms 'reduction' and 'micro-reduction' to refer to reductions as described in this section. So from now on 'reduction' does not imply an ontological identity relation, but an explanation or description of macro properties by properties of its parts.

3 Emergentism

3.1 Three kinds of emergence

Bedau distinguished three kinds of emergence (Bedau 2002), this distinction provides a framework in which most views of emergence can be placed. As mentioned in chapter 1, Bedau uses two criteria as a mark of emergence. The first criterion holds that emergent phenomena depend on underlying processes, the second states that emergent phenomena are autonomous from their underlying processes. According to Bedau, the three forms he distinguishes all satisfy these two conditions.

3.1.1 Nominal emergence

The first kind of emergence Bedau describes is nominal emergence. A nominal emergent property cannot be a micro property. Nominal emergence does not explain the origin of micro and macro properties, they are just assumed to be identifiable and the macro properties are called nominal emergent. Bedau gives two examples of nominal emergent properties (Bedau 2002). The first example is the shape of a circle. Circles consist of collections of points and these individual points have no shape. Together all the points have the property of being a circle and the single points do not possess this macro property. The second example he provides is the particular weight of a pile of stones (Bedau 2002). If all the separate stones weigh one kilo, twenty stones together weigh twenty kilo. The macro property of weighing twenty kilo is not possessed by one single stone, only by the pile of stones. A third property, also nominal emergent according to me, is being a pair. To be a pair, two entities are needed: two shoes are a pair of shoes, one single shoe is not and two persons can be a pair, but one person cannot.

According to Bedau, nominal emergence satisfies his two constraints of emergence. Nominal emergent macro properties are dependent on the micro processes because the whole consists of the parts. At the same time the macro properties are autonomous from the micro processes, because the macro properties do not apply to the micro parts.

The notion of nominal emergence is very broad, it covers a lot of phenomena. But exactly this broadness of nominal emergence is its weak point, it applies to *all* macro phenomena and not just to special cases. Smart, an identity theorist, wrote the following.

“Of course the physicalist will not deny the harmless sense of “emergence” in which an apparatus is not just a jumble of its parts.” (Smart 1981)

Nominal emergence would classify the phenomena, which Smart refers to as emergent, but most emergentists will not consider this as genuine emergence. Bedau (2002) rejects his two examples above as being interesting emergent properties, because the macro properties are easy to derive and predict. All the points in the figure have the same distance to a given point (so the figure must be a circle) and twenty times one kilo is twenty kilos. I think that the three examples above are not really interesting, because the parts in the examples do not interact with each other. The nominal emergent macro properties are only descriptions of the parts together, but they do not say anything about new phenomena that arise because of the interactions of the parts. I propose to make a distinction between macro properties with and without the involvement of interactions of the parts. I will refer to macro properties without interactions of parts with the term *static emergence* and to macro properties involving interactions with *process emergence*.

I chose the three examples of phenomena with very simple macro properties to show a specific feature of nominal emergence. Only nominal emergence would classify these examples as emergent, other kinds of emergence would classify them as non-emergent properties. Other macro properties, like life and intelligence (process emergent in my terms), cannot be predicted and explained as easy as the examples above. According to most emergentists, these are the real emergent properties. So nominal emergence includes both static and process emergent phenomena, most other kinds of emergentism only cover process emergent phenomena. In this thesis I want to focus on process emergence and to pick out only these special, more complicated macro phenomena, the notion of emergence has to be more restricted.

3.1.2 Strong emergence

A stricter notion of emergence Bedau distinguishes is strong emergence (Bedau 2002). This notion adds the requirement to emergent properties to be *supervenient properties with irreducible causal powers* (Bedau 2002). In this thesis the term ‘strong emergence’ will only refer to Bedau’s account of it. The supervenience principle says the following.

Let A and B be two sets of properties: then A supervenes on B if and only if necessarily any two things that have the same properties in B have the same properties in A (that is, B-indiscernibility entails A-indiscernibility). (Kim, cited in Beckermann 1992)

In the mind-body debate the principle is sometimes put as ‘no mental difference without a physical difference’ (Kim 1996). The concept of supervenience is often connected with dependence or determination. Then the supervenience principle is read as if a macro level (with the properties of set A) is totally determined by a micro level (with the properties of set B). But this is not what the supervenience principle says, it only makes a claim about how two sets of properties covary with each other (Kim 1996).

Besides the supervenience of the macro properties, strong emergent properties do also have irreducible causal powers, with effects at both the macro and the micro level (Bedau 2002). The phenomenon of effects from the macro to the micro level is termed *downward causation* (Bedau 2002). Only macro properties are able to cause these downward effects, they cannot be caused on the micro level itself. The new macro properties cannot be explained by or derived from the properties of the components, their causal influence does not occur via the activity of the micro-properties which constitute them (Bedau 1997). Strong emergentists as Bedau describes them, claim that if a system with highly related components reaches a certain level of complexity, genuinely new properties with these macro causal powers arise.

The notion of strong emergence satisfies Bedau’s two conditions of emergence. The macro properties are supervenient properties, supervenience explains how the emergent properties depend on their underlying bases. The second criterion, the autonomy of the macro properties from underlying processes, is satisfied because of the irreducibility of the macro causal powers.

Strong emergence resembles the first notions of emergence, as they originally arose in the middle of the nineteenth century (McLaughlin 1992). The very first emergentists of this British Emergentism, all adhered to more or less the same version of emergence. Their basic assumptions were: everything is made of matter, nothing happens without some motion of elementary particles and all motion behaves according to the laws of mechanics. And their emergentism, like strong emergence, was also committed to downward causation (McLaughlin 1992). They thought that certain types of structures entailed emergent causal powers. The laws according to which the emergent causal powers arose, were not ‘reducible to’ or ‘derivative from’ lower level laws. These ideas exactly cover the notion of strong emergence.

British Emergentism reached its top in the first quarter of the twentieth century and it was very popular at that time. But after its rise and a successful period, its problems were emphasized and it lost popularity.

Kim (1996) emphasized a problem with downward causation, called the *exclusion argument*¹. Strong emergence involves downward causation, macro properties having causal effects on the micro level and at the same time supervening on micro parts (Bedau 2002). All the upward arrows in the figure below mean the supervenience of an emergent property E or E* to a base property B or B*. Specific micro conditions expressed by B or B*, always lead to corresponding macro conditions E or E*. The first picture of the figure shows an emergent macro property E, which has causal effects on the macro level and causes a new emergent property E*. The second picture shows a macro property E having causal effects on the micro level, E causes (by downward causation) a new base property B*. The conditions of B* realize a new emergent property E*. Strong emergence claims same-level (picture 1) and downward causation (picture 2) to be possible (Bedau 2002).

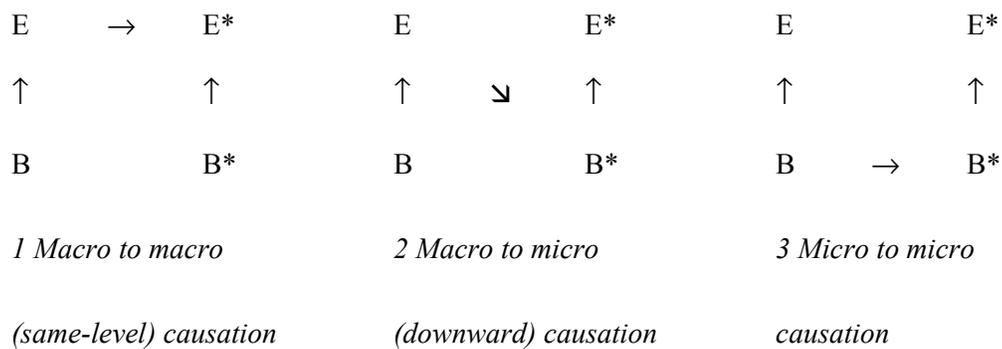


Figure 2: Different types of causation (Kim 1996, 230-232)

Strong emergentists also agree with micro level influence of the parts of a system upon each other and the supervenience of emergent properties on these micro level configurations, shown in the third picture of figure 2 (Bedau 2002). The problem Kim (1996) denotes is to unite the causation types of the first two pictures with causation just on the micro level.

The problem with the union of picture one and picture three is the following. The supervenience principle states that equal micro conditions always appear with the same macro conditions. This implies that different macro conditions presuppose different micro conditions. So if E causes a different macro property E* (picture one), B has to be different from B*. This change on the micro level has to obey the physical laws, so B has to cause a B*. This is shown in picture three and the result is a new B* realizing a new E*. Strong emergentism claims that the causation from E to E* has extra causal powers, causing more than the micro causal processes

¹ In his book *Mind in a physical world* (1998) Kim adjusts his argument: the argument only applies to within level relations, for micro-macro relations Kim acknowledges ‘new causal powers’. So the argument as presented here is not up to date, but the discussion helps to clarify the concept of downward causation.

and not explainable by the micro causal processes (Bedau 2002). According to strong emergentists, an example of these macro causal powers are mental phenomena: they say that mental causal powers are not explainable by activities of the neurons (Bedau 2002). But how are these extra causal powers possible, without ignoring the physical laws on the micro level?

The union of pictures two and three of figure 2 is also problematic. All the micro processes obey to the laws of physics and cause the next micro situation. This is shown in picture three: base property B causes a new base property B* and the base properties both realize an emergent property. No place seems to be left for extra macro causal powers on the micro level. In picture three causation from E to B* (shown in picture two) is not necessary to form a new emergent property E*, because B* and thereby E* is already caused by B. What extra could E cause on the micro level, which cannot be caused by B?

Kim believes the third picture to be the right representation of reality. Because this picture is not compatible with macro to macro and macro to micro causation, Kim rejects views like strong emergence including macro causation with extra causal powers. I do not totally agree with him. The real problem is not the extra causal powers of emergent properties per se, but the connection strong emergentists make between the macro and the micro properties. A property of a system can be described by the causal powers it has. So a system with an emergent property its parts do not have, also has certain causal powers as a whole its parts do not have. But instead of talking about a whole acting on its parts, it is better to say that macro causality has micro causal implications because the macro level is made up from the micro constituents. Macro properties are relevant for the micro level, but they do not have causal effects acting against the processes of the parts.

As I understand Bedau's notion of strong emergence, strong emergentists state that the macro level cannot be derived from the micro level, even with complete knowledge of it. So no explanation of macro causal powers by micro properties is possible. Without such a connection between the micro and macro level, Kim's pictures form a problematic obstacle. Then a conflict arises between micro and macro powers about which of the two is the real causer of new phenomena. The extra macro causal powers then seem to have more powers than for example the laws of motion and collisions, as if they have vital or magic powers. This is not very plausible, therefore the macro causal powers have to be in some way derived from or explained by the properties and interactions of the micro level parts. Total disconnection of the micro and the macro level as in strong emergence should be avoided.

3.1.3 Weak emergence

The last notion of emergence, weak emergence, is an intermediate position between nominal and strong emergence. Weak emergence refers to the aggregate global behavior of certain systems, in which the weak emergent macro behavior is the total of all the micro processes and their interactions together. The complexity of these macro processes yields the impossibility of giving simple explanations for them. According to Bedau (1997), the central idea of weak emergence is that emergent causal powers can be derived from micro-level information, but only by simulation. He thinks the only way to predict what will happen in an emergent process, is by executing all the steps of that process. If different ways than simulation exist to predict the outcome of a process, then the process is not emergent.

Later I will discuss Bedau's criterion of 'prediction by simulation only' and consider its usefulness, for now it is easier to outline weak emergence by marking what it is not. Weak emergence is not strong emergence, because within weak emergence the causal powers of the emergent can be explained from properties of the parts on the micro level. Strong emergentists say this is not possible. Weak emergence is different from nominal emergence, in the sense that nominal emergence is broader and classifies more phenomena as emergent. Bedau's definition of weak emergence implies that emergent phenomena must be processes. Static emergent phenomena are predictable by other ways than simulation, they do not change and no simulation is needed to predict their future. So weak emergence classifies the static macro phenomena, emergent according to nominal emergence, as non-emergent.

Kim's criticism to strong emergence (the exclusion argument) also applies to the notion of weak emergence. Weak emergence also involves downward causation and new macro causal powers (Bedau 2002). The problems of strong emergence with the causation types in the pictures of figure 2, also apply to weak emergence. But as I argued, downward causation per se is not problematic. An example of downward causation is the automotive power of a car. The individual car parts do not have the power to move, but a car as a whole has. When a car is moving, the macro property of the car causes the (micro) parts to move. This example shows what is meant with downward causation in weak emergence and that it is not problematic. Because we know how cars work, it is possible to explain how the car parts and their interactions create the macro property of automotive power. It is clear that this macro power also has (downward) implications on the micro level.

In other examples of downward causation, like mental causation, we do not totally understand how mental properties arise out of micro parts. At present it is not possible to exactly

explain, derive or predict the macro level from the micro level in this context. Strong emergentists see this as the characteristic feature of emergent properties: emergent properties are not explainable, derivable or predictable from information about the micro level in any way (Bedau 2002). According to weak emergentists, it is possible to predict the macro level from knowledge of the micro level, namely by simulation (Bedau 2002). They will agree that we cannot (yet) predict all mental phenomena from information of the micro level, but according to them this is because we do not have complete knowledge of them micro level. If we knew more about the micro level, we would be able to predict the macro processes, by simulating the micro processes.

Strong emergentists call the new, strong emergent macro powers, unexplainable and unpredictable from properties of the micro parts, *irreducible* causal powers (Bedau 2002). Bedau (2002) does not use the term 'irreducible' for weak emergent macro powers. It depends on the meaning of irreducibility whether weak emergent macro powers are reducible or not. They are not reducible in the sense that macro phenomena have an ontological reductive link with micro phenomena and that they are ontologically identified with them. But weak emergent causal powers are reducible from the micro level in the sense that there is a link between the micro and the macro level. It is possible to explain or predict macro properties from micro properties.

The weak emergence Bedau defends seems to be a right notion of emergence to proceed with in this thesis. The arguments against emergence including irreducible (in the sense of unexplainable) causal powers are strong, the total independence of the macro level from the micro level as in strong emergence seems hard to combine with the supervenience of the macro level on the micro level. For this reason I reject strong emergence from further consideration here. Then nominal and weak emergence are left. Nominal emergence is very broad and includes both static and process emergence, weak emergence is a subset of it and only includes process emergence. In this thesis I want to focus on process emergent phenomena, so weak emergence picks out the more interesting cases of nominal emergence and leaves the simple macro properties outside.

In this thesis weak emergence is a suitable notion of emergence, but it is still a vague concept. In the remainder of this chapter I will go deeper into the subject and describe more specific conditions to decide when to speak of emergence.

3.2 What makes something emergent?

From now on the term emergence will refer to weak emergence, unless otherwise stated. Up to this moment the meaning of emergence in this thesis is not more specific than something that is weaker than strong emergence, but more specific than just nominal emergence. In this section more explicit criteria are provided to distinguish between the emergent and the non-emergent. Andy Clark said the following about emergence.

“We need to find an account of emergence that is neither so liberal as to allow just about everything to count as an instance of emergence, nor so strict as to effectively rule out any phenomenon that can be given a scientific explanation.” (Clark 2001, p 113)

What Clark means with too liberal corresponds more or less with Bedau’s notion of nominal emergence. What he means with too strict is that not everything with a scientific explanation is reckoned as a simple uninteresting macro property. After this statement Clark sums up the following suggestions to divide between the emergent and the non-emergent (Clark 2001).

- 1 *Emergence as collective self-organization.* To this account, an emergent phenomenon is any interesting behavior that arises as a direct result of multiple, self-organizing (via positive feedback and circular causation) interactions occurring in a system of simple elements.
- 2 *Emergence as unprogrammed functionality.* Emergence here is a kind of side effect of some iterated sequences of agent-world interactions. The behaviors can only be indirectly manipulated.
- 3 *Emergence as interactive complexity.* Here emergence means effects, patterns or capacities made available by a certain class of complex interactions between systemic components.
- 4 *Emergence as uncompressible unfolding.* This last suggestion holds that a phenomenon is emergent if its prediction requires simulation.

Two main directions can be recognized in these four suggestions of categorizing emergence. Suggestion number four is one way to look at emergence, I will discuss this idea of *no prediction without simulation* in the next sub-section. The second direction is the third option, because this suggestion is a combination of number one and two. I will discuss the notion of *emergence as interactive complexity* in subsection 3.2.2.

3.2.1 Prediction by simulation only

As already mentioned, Bedau thinks that something is emergent if the only way to predict what will happen is by simulation. He defines weak emergence as follows.

“Macrostate P of S with microdynamic D is weakly emergent iff P can be derived from D and S’s external conditions but only by simulation.” (Bedau 1997)

This definition can be illustrated by applying it to the example of emergent flocking behavior of birds. The system S in this example is a flock of birds. The macrostate P is the flocking behavior, microdynamic D are the birds and the rules according to which they behave. If all the information about microdynamic D (the position, velocity and direction of all the birds and the rules) and system S’s external conditions (open air or obstacles) is available, macrostate P (the position of the birds after a certain amount of time) can be predicted. According to Bedau (1997), no algorithm different from calculating or executing all the steps (movements) can predict this macro state P, simulation is the only way.

Not everybody agrees with the idea that prediction by simulation only is *the* discriminating factor of emergence. Clark (2001) for example thinks that the condition is overly restrictive, it is often sufficient to simulate only a subset of actual interactions for predicting how the patterns will form and unfold over time. He says that Bedau unjustly restricts emergence to phenomena that resist all such attempts at low-dimensional modeling. So whereas Bedau stresses the unpredictability of emergent phenomena, Clark, and also Daniel Dennett, think that predictable patterns often are emergent patterns. Dennett (2003) says that predictability is characteristic of emergent patterns, according to him total chaos (randomness) is not emergent. Dennett stresses that although something is determined on a micro level, new patterns at higher levels do arise. These patterns are real and more than the sum of the parts. If the higher-level phenomena were not predictable at all, there would not be macro patterns.

Some confusion is going on in this discussion. The predictability Dennett and Clark are talking about, is not the same predictability as Bedau means. When Dennett takes predictability as a characteristic of emergence, he is talking about the prediction of global patterns. But Bedau would probably also agree that emergent phenomena are often phenomena with returning recognizable patterns. Bedau is talking about the prediction of exact configurations instead of the global patterns, and these are more difficult to predict. Although the global patterns of a particular emergent phenomenon might be predictable, the very specific details might not. So if the criterion of ‘predictable by simulation only’ is used, the detailed predictability of a specific configuration is meant.

3.2.2 Complexity

Of the four kinds he suggested, Clark denotes complexity as the best option to classify emergence. He says the following about it.

*“Emergent phenomena as the effects, patterns, or capacities made available by a certain class of **complex** interactions between systemic components.” (Clark 2001, p 114)*

Other philosophers, who advance the same position about emergence and complexity, are Bechtel and Richardson. In their mechanistic approach they claim to take a middle position. They want to remain physicalists (there is no special vital power or mind), but at the same time not endorse a reductionistic mechanism. In a mechanistic perspective it seems entirely natural to recognize higher levels of organization as semi-autonomous, but also as regulating behavior of components at lower levels. About emergence they say that

*“Emergence is a consequence of **complex** interaction.” (Bechtel & Richardson 1992, p 285)*

So also according to them emergence has to do with complexity, but what does that mean?

Complexity can be found in many different kinds of systems, from computational systems of pure information to biological systems such as cells and organisms to natural ecosystems and human society (Salen & Zimmerman 2004). Jeremy Campbell studies complexity by looking at

“self-organizing, replicating, learning and adaptive systems.” (Salen & Zimmerman 2004)

Chris Langton, a mathematician and inventor of the notion of artificial life, differentiates four kinds of systems (Salen and Zimmerman 2004), each category refers to a different degree of complexity. The order of the four kinds below is from less to more complex systems.

- 1 *Fixed systems.* These systems remain unchanged and the relationships between the elements are always the same.
- 2 *Periodic systems.* These are simple systems, which endlessly repeat the same patterns.
- 3 *Complex systems.* These systems are more complicated and unpredictable than a periodic system, but not as full of dynamic relationships as totally chaotic systems.

- 4 *Chaotic systems.* The elements in a chaotic system are constantly in motion and their states and relationships are random.

How can this help to say something about emergence? It seems obvious that emergent phenomena at least arise in the third category of complexity. But what is the exact definition of this category? Most people will probably also agree that macro properties in the first category are non-emergent macro properties. Only nominal emergentists would probably describe emergent properties to these static, unchanging systems. But what about the periodic systems, can they create genuine emergent phenomena? But even if emergent phenomena only occur in complex systems, it is still a question whether *all* complex processes are emergent processes, and if not, which part of the complex processes contains emergent properties.

3.2.3 Conclusions

Complexity and prediction by simulation often apply to the same situations. Some people include 'prediction by simulation only' as part of the definition of what it is to be a complex adaptive system in general (Holland 1992, cited in Bedau 1997). Whether that is right or not, both conditions are not identical. Right now it is important to determine their values as indicators of emergence. That is not easy, complexity and prediction by simulation only both seem to have a lot to do with emergence but it is still not clear when to speak of emergence and when not.

To my opinion complexity as a criterion of emergence has some advantages over prediction by simulation only. First, complexity allows properties to be more or less emergent. A process can be more or less complex, so it also makes a degree of emergence possible. When prediction by simulation only is used as a criterion for emergence, it would be an all or nothing affair. A process is predictable by simulation only or it is not, there is nothing in between. In the second place, it is difficult to decide when something is only predictable by simulation or when other algorithms could also do the job. It is possible that at this moment a process is only predictable by simulation, but only because the algorithm predicting the behavior has not been found yet. Then according to Bedau the process would not be emergent, but this would not be known. In the third place, if predictability by simulation only refers to the exact configurations of the micro parts, it could also apply to non-emergent phenomena. The exact locations of molecules in a gas for example: this might be predictable by simulation only, without the gas having emergent properties. The final reason is a practical one: simulation would not be workable because in most cases it is impossible to completely simulate a certain process.

Because of these three reasons I conclude that prediction by simulation only is not *the* marking criterion of emergence, but it does provide the connection between the micro and the macro level missing in strong emergence. I think that in theory macro phenomena can always be predicted by simulation with complete knowledge of the micro level, maybe sometimes there are other ways to predict the continuation. So a micro and a macro level are at least connected with each other because of the predictability of a macro phenomenon by simulation of the micro level processes. The complexity criterion of emergence seems useful to me, but complexity should be better defined and the exact relation between complexity and emergence should be explicated.

3.3 What makes a property non-emergent?

So far I described different forms of emergence and different ways to recognize emergence. The conditions of emergence I gave were all positive determinants of emergence, their presence indicated emergence. In this section the approach will be different, conditions of non-emergence will be given. When a property fails to meet such a condition, the property is emergent. This section about non-emergence will be based on William Wimsatt's (1986) survey of conditions of aggregativity.

3.3.1 Forms of aggregativity

Wimsatt (1986, 1997) analyzed the ways a system can be put together, he called these different ways forms of aggregativity. According to him it did not make sense to speak of a total system being or not being an aggregate of its parts (Wimsatt 1986). A physical system always exists of a set of parts, but sometimes the properties of a system are just aggregates of their parts and sometimes they are not. An example of a system property merely being an aggregate of its parts is the weight of a pile of stones, which is nothing more than the sum of the weights of the parts (Wimsatt 1986). A living organism has properties as a whole, which its parts do not possess. Here the whole is more than the sum of the parts and not just an aggregate of them (Wimsatt 1986). So Wimsatt approaches aggregativity as a kind of relation between properties of systems and properties of their parts.

According to Wimsatt (1997), aggregativity is the state in which 'the whole is nothing more than the sum of the parts'. When a system property is just a mere aggregate of its parts, it is not emergent. An emergent property is a system property, which is dependent upon the mode of organization of the system's parts. Wimsatt thinks the easiest way to distinguish whether a

property is emergent or not, is defining when it is not by summarizing the different forms of aggregativity (non-emergence). Therefore, he looks at different ways a system can be put together in which the property of the whole does not depend on specific organizations and properties of the parts. The less a property of a whole depends on specific organizations of the parts, the higher the level of aggregativity and thereby non-emergence. He discerns four forms of aggregativity and if one of these four conditions is not met, the system is not just a mere aggregate of its parts and we can already speak of emergence. Below Wimsatt's four conditions of aggregativity in a system are shown.

- 1 Intersubstitutivity of parts
- 2 Qualitative similarity with a change in the number of parts
- 3 Stability under reaggregation of parts
- 4 Minimal interactions among parts

(Wimsatt, cited in Bechtel 2003)

Wimsatt's *first* condition deals with whether a part can be intersubstituted without affecting the system property. The higher the change of the system property when parts are substituted, the lower the level of aggregativity. An example that fails to meet this condition is a string of genes. The location of the genes is very important to the properties of the whole, the genes cannot be intersubstituted without consequences.

The *second* condition of aggregativity is about the similarity of system properties under addition or removal of parts from the system. If addition or removal of parts does not affect the system, it has a higher level of aggregativity. An example of a property that does not satisfy this condition is the stability of a round arch, if the arch is only stable with a specific amount of stones. This condition also rules out a variety of threshold phenomena as aggregative properties.

The *third* condition is about invariance of a system under operations involving decomposition or reaggregation of its parts. When a system is unaffected by redistribution of its material into different parts, it answers to this aggregation condition. Examples of this kind of operations are recombination, transposition, inversion and translocation of blocks of genes on a chromosome. If these operations for example on genes organized into a chromosome have effects on the properties of the chromosome as a whole, the properties of the chromosome fail to meet this third condition of aggregativity.

The *fourth* and last condition deals with cooperation and inhibition. A system has a higher level of aggregativity when there is no mutual cooperativity or inhibition among the parts

of a system. Examples of failures to this condition are atoms in a self-amplifying process, which cause bombs to explode under certain circumstances.

The properties expressed by genes, the stability of an arch, the properties of a chromosome and the power to let something explode all fail to meet at least one of the four conditions of aggregativity. This means that all four properties are emergent properties. According to Wimsatt, emergence is not something exceptional, on the contrary, it can be found everywhere around us. A lot of or even most system properties are emergent in some way.

Wimsatt's model provides a lot of insight in the discussion about emergence. Analyzing the way a system is built up is a useful way to gain more insight about emergence and to my opinion Wimsatt is right in attributing much importance to the properties of a system. In this thesis however, I want to pay attention to macro phenomena involving interactions among the parts. But as within nominal emergence, Wimsatt's model also uses a broad definition of the concept of emergence and classifies static macro phenomena as emergent too. An example of a static nominal emergent property I gave in the subsection about nominal emergence was 'being a pair'. This macro phenomenon is not a process and would not be classified as emergent by weak emergentism, the kind of emergentism I adopted in this thesis. But being a pair fails to meet Wimsatt's second criterion of aggregativity, the property is not similar with a change in the number of the parts. So according to Wimsatt, being a pair is an emergent property.

Wimsatt's account of emergence and the notion of nominal emergence neither one require interactions among different parts as a condition of emergence. This makes the two accounts of emergence too broad for this thesis and interactions among different parts should be a condition of emergence here. The requirement for an emergent phenomenon to be a process is already present in the two other criteria of emergence, 'prediction by simulation only' and 'complexity'. Simulation implies the execution of steps and in a stable, unchanging system no steps are taken. According to the complexity criterion, unchanging systems are classified as fixed systems instead of complex systems.

So to obtain a workable notion of emergence for this thesis I suggest to restrict Wimsatt's model of emergence. Therefore, besides his own conditions of emergence, the condition that changing relations between parts have to play a role in emergent phenomena should be added. This makes emergence apply to dynamic phenomena that fail to meet at least one of Wimsatt's four conditions. Macro properties of emergent phenomena can be explained by properties on the micro level, at least by simulating the micro processes.

4 Case study

The case study in this chapter is the study of a computer simulation in which a micro and a macro level can be recognized. After a description of the simulation case in the first section, a reductionistic and an emergentistic account of the relation between its micro and macro level will be given.

4.1 A description of the simulation

4.1.1 Cellular automata

In the first chapter, the Game of Life (cells on a grid which together form gliders, eaters and shooters) was introduced as an example of a process with different levels of organization. I recall this example, because the Game of Life makes use of *cellular automata* and the simulation case discussed in this chapter also involves cellular automata. The basic features of cellular automata are the following.

- Cells are arranged in a regular D-dimensional grid
- Every cell adopts a state from a finite set of states
- Time is discrete
- Cells change their states according to local rules
- The same transition rule applies to all cells
- In each period cells are updated

(Hegselmann & Flache 1998)

A lot of real processes can be simulated with cellular automata, from the spreading of viruses to the behavior of fishes in a shoal to social human behavior. In all the simulations a set of comparable parts (viruses or fishes or humans) exhibit the same behavior. In simulations the starting conditions can easily be changed and effects of changes on the micro level to the global behavior can quickly be obtained. This is a big advantage of simulations over real experiments and simulations with cellular automata have often turned out to yield surprising and very useful results.

4.1.2 The simulation

The simulation case discussed in this chapter is a model of social dynamics. J.M. Sakoda designed the first version and his goal was to better understand group formation (Sakoda 1971). Sakoda used an eight times eight ‘checkerboard’ for his simulation, with two groups of six members ‘living’ on it. A random process determined the starting positions of the members, initially the individuals were always randomly distributed over the board. The individuals on the board had a positive, a neutral or a negative attitude towards their own group and to members of the other group. The attitudes of group x and group y are expressed in a preference table.

	Towards own group	Towards other group
Attitude of x's	+1, 0 or -1	+1, 0 or -1
Attitude of y's	+1, 0 or -1	+1, 0 or -1

Figure 3: A general preference table (Sakoda 1971)

Individuals could move to the eight cells around them if they were empty. If there were no empty cells then an individual was allowed to jump over one individual, but no further than that. The members of each group made their moves in a random order. During a cycle each individual got one turn. The valuation formula below determined whether and to which direction the individuals moved. An individual moved to the cell with the highest outcome of the valuation formula, if the value was highest in its current cell it would not move.

$$\sum_{j \in P} \frac{V_{ij}}{\sqrt{d_{ij}^2}} \mathbf{w}$$

Formula 1: Valuation formula (Sakoda 1971)

Here **P** expresses the set of all individuals, V_{ij} is the valence of the attitude of individual *i* to individual *j*, d_{ij} is the Euclidean distance of individual *i* to individual *j* and **w** is the distance weight. Individuals towards which a particular individual had a positive attitude attracted and a negative attitude made the individual to avoid them. The smaller the distance between individuals was, the stronger their influences on each other. The value of **w** determined how strongly valences were discounted by distance: the greater **w**, the less the valences were discounted by distance. In all his examples, Sakoda used $\mathbf{w} = 4$ in the experiments (Sakoda 1971).

With this model, Sakoda (1971) discovered that individuals with positive attitudes to each other and negative attitudes to the members of the other group quickly formed two groups. He called this condition *segregation*. Further he discovered that individuals with neutral attitudes to each other also formed a segregated group if they had negative attitudes to members of the other group. He called this condition *suspicion*. The preference tables below show the attitudes used in the experiment.

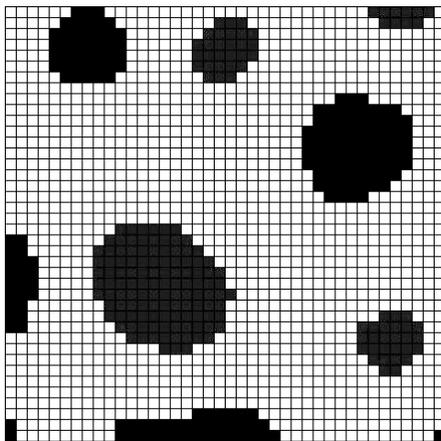
Segregation	Black	Gray
Black	1	-1
Gray	-1	1

Suspicion	Black	Gray
Black	0	-1
Gray	-1	0

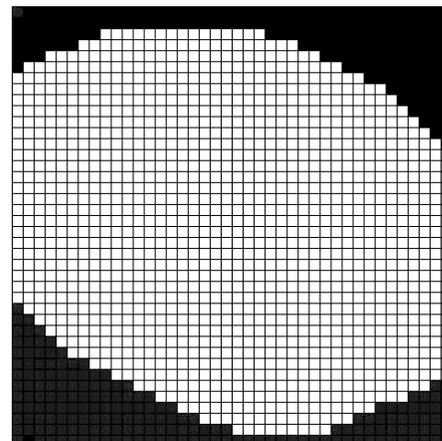
Figure 4: The preference tables of the segregation and the suspicion condition (Sakoda 1971)

What Sakoda did not see was that a negative attitude towards the other group combined with indifference towards one's own group led to much more massive clusters than in combination with positive feelings to one's own group.

This effect was discovered by Hegselmann and Flache (1998) when they replicated Sakoda's simulation on a bigger field. Hegselmann and Flache used a grid of 40 times 40 cells, with two groups of 180 members living on it. Besides this, everything else was kept constant. Every run of the simulation with a random initial distribution of the individuals over the cells resulted in the same kind of patterns. A random distribution here is a random spread of black and gray individuals over the total field. Two possible outcomes of a run of the simulation are shown in figure 5.



Segregation



Suspicion

Figure 5: Outcomes of the segregation and suspicion condition (Hegselmann and Flache 1998)

In the first figure a pattern of islands with members of the same group has arisen. Despite the neutral instead of positive attitudes towards the own group, more separated clusters and bigger formations of members of the same group have been formed in the second condition.

4.1.3 Macro and micro levels in the simulation

At first sight the outcomes of the simulation case might not seem very obvious: the clusters of individuals of the same group are bigger in the condition without positive attitudes to members of the same group. But with the results in the figures it is easy to explain why these patterns originate. In both conditions the individuals move away from members of the other group. But in the first condition the positive attitude towards the own group influences this process of ‘running away’, in the second condition nothing stops this process. Although many different end-situations are possible, the final global patterns are the same each run. The differences between runs take place on the micro level, the states of the separate cells are nearly always different. The general features remaining constant during different runs are macro phenomena.

The description of the simulation case exactly explains the micro level processes and properties. The specification of the macro patterns is more difficult, till now I only mentioned them ‘returning patterns on the macro level’. But what are these returning and recognizable phenomena in every run? To show the difficulty of a good account, I made a start with the formulation of a definition of the final macro pattern in the segregation condition. The most striking feature of these patterns is the organization of islands with individuals. A first approximation of the definition of such an island could be this:

“A stable, more or less round area of adjacent cells, totally occupied by individuals of the same group, surrounded by empty cells”.

This might seem to correspond with an island on the grid, but the definition leaves much room for ambiguity. For example: what does ‘more or less round’ exactly mean? How small can an island be, could one single cell also be an island? How big can an island be, is an island including all the members of one group allowed? Islands on the borders of a grid are not round, what about them? Besides all these problems, much more than the definition of an island is needed for the definition of the whole pattern. The definition should take into account: the number of islands that appear in a pattern, the distances between the islands, the distribution of the islands of the two different groups over the grid, etcetera. The example of the segregation condition obviously does not

provide a complete definition of the whole macro pattern, and my attempt shows that this the task is not very simple.

The final macro pattern of the second condition, suspicion, is somewhat easier to define. The following definition neither will be *the* totally complete and correct definition, taking all aspects into account. However, the definition will be useful in the further analysis of the simulation case. In the suspicion condition most individuals end in a location as far away as possible from individuals of the other group. Small exceptions are possible, black individuals can for example imprison a gray individual. But with the conditions of Hegselmann and Flache, in general the sum of the distances between all the black and all the gray individuals is maximized in the end situation. That is, the course of running the simulation results in a relative maximum of the sum of the distances, no allowed single move increases this value. The calculation of the sum of the distances is shown in the formula below.

$$\sum_{i \in G} \sum_{j \in B} (d_{ij})$$

Formula 2: Sum of the distances

Here **G** is the set of gray individuals (**i**'s) and **B** is the set of black individuals (**j**'s). **d_{ij}** is the Euclidian distance between individual **i** and individual **j**.

4.2 Reductionistic approach

Two possible ways to describe the simulation case are a description of the micro level phenomena and a description of the macro level phenomena. Reductionists say that both descriptions refer to the processes in the simulation. They try to find the connection between the two ways of describing the simulation, by reducing the macro description to the micro description via deriving the former from the latter. In this case an obvious part-whole relationship brings about the differences between the descriptions. Therefore, as treated in chapter 2, micro-reduction is the best way to describe the relation between the two levels here. The reduction will show that every action in the simulation case brings the distribution of the individuals over the cells closer to a final distribution with a certain macro pattern.

4.2.1 A reduction of the suspicion condition

To give a reduction of the simulation case, one could show why every step in the simulation is a step in the direction of a certain end-situation with a particular macro pattern. In section one, I defined the final outcome of the suspicion condition as a relative maximum of the sum of the distances between all the individuals of the different groups. So to reduce the suspicion condition it has to be proved that the sum of the distances will never decrease during a run of the simulation in the suspicion condition.

There are two possible situations in one turn of a simulation run: an individual moves or an individual does not move. If the outcome of Sakoda's valuation is best in the individual's current position, the individual will not move. Nothing changes on the grid and the sum of the distances remains the same. If the individual does move and you do not want the sum of the distances to decrease, then for $w = 4$ the following condition should hold with d'_{ij} indicating the distances after the move:

$$\text{If } \sum_{j \in B} (1/\sqrt{d'_{ij}}) < \sum_{j \in B} (1/\sqrt{d_{ij}}) \text{ then } \sum_{j \in B} (d_{ij}) < \sum_{j \in B} (d'_{ij})$$

The if-part of the condition represents Sakoda's valuation function. If a move satisfies this condition, then the individual indeed will move and according to the second part of the condition the sum of the distances should increase. In most cases an individual's move will satisfy this condition, but exceptions are possible. However, the condition does hold for moves in which none of the distances decreases. This idea can be used to add an extra condition which makes it very plausible that the sum of the distances will not decrease during a run of the simulation case.

In the situation as it was, the move with the highest value according to Sakoda's valuation function (formula 1, p 41) was picked. In the new situation with the extra condition, the distances between individuals are also considered. First all possible moves are determined according to Sakoda's valuation function. Then, from the possible moves in which none of the distances decreases, the one with the highest value according to Sakoda's valuation function is picked. If there are no such moves, just the move with the highest value according Sakoda's function is picked. So if it is not necessary, none of the distances decrease and in these situations the above criterion automatically holds. Note that this extra condition has some plausibility in the light of the suspicion condition: if possible do not come closer to any individual of the other group. Although this extra condition does not totally exclude the rise of some sort of loops, it may

be expected that in most cases the same macro pattern as in the old situation will arise. This shows that ‘relative maximum of the sum of the distances’ is a satisfying description of the macro pattern on the micro level.

4.2.2 Concept reduction versus theory reduction

Chapter 2 ended with a section about micro-reduction, the form of concept reduction in which aggregate concepts played an important role. One of the conclusions of the section implied that aggregate concepts in micro-reductions should be about macro properties of a system instead of the systems themselves. The definition of an aggregate concept was a concept defined in terms of micro types of micro objects for aggregates of such micro objects. An aggregate concept used in the reduction of the simulation case is the sum of the distances between all the black and gray individuals. This is indeed a definition of a macro concept in terms of micro types of micro objects. The micro objects are the black and the gray individuals, and aggregates of them can form macro patterns on the grid. A micro type relation of the micro objects is the distance between two individuals. So the requirements of a formulation of an aggregate concept seem to be fulfilled in the reduction of the simulation case, but what about the subject of reduction?

My criticism of the description of ‘water’ as ‘an aggregate of H₂O molecules’ was the neglect of the macro properties of water. I stated that aggregate concepts should refer to properties of systems and not to systems themselves. In the reduction of the suspicion condition the aggregate concept ‘sum of the distances’ is a state concept, which only refers to a part of the characteristic macro phenomena in the simulation case. But a little change in the formulation of the aggregate concept does make it apply to an important property of the simulation case, the course to the specific macro pattern of the suspicion condition. If the formulation of the aggregate concept is changed to ‘maximizing the sum of the distances’, then it is a process concept, which refers to the actions during the simulation case. Formulated this way, more attention is paid to the macro properties in this reduction than for example in the reduction of water.

Till now I only discussed concept reduction. The formulation of the most important concept (maximizing the sum of the distances) however refers to a regularity in the simulation case and regularities are a kind of laws. So although a goal of this thesis was to apply concept reduction to the simulation case, maybe its analysis could also involve the reduction of laws. According to Kuipers, applying his model of theory reduction, the reduction of the simulation case would contain an application step, an aggregation step and an identification step. This would deliver the following reduction (Kuipers, personal communication).

- 1 *Application step*: The refined condition for moving makes it increasingly probable for any $i \in G$, respectively $j \in B$, that $\sum_{j \in B} (d_{ij})$, respectively $\sum_{i \in G} (d_{ij})$, will not decrease.
- 2 *Aggregation step*: This step results in the aggregated law that a random start almost always leads to a relative maximum of the sum of the distances.
- 3 *Identification step*: By identifying the sum of the distances with the distance between the two groups, the aggregated law amounts to the macro law that random initial conditions in the simulation always lead to a macro pattern with two separate, stable groups on a relative maximum distance of each other.

4.2.3 Conclusions

The reduction in section 4.2.1 explains how all micro level actions contribute to the formation of the macro level pattern. So a derivation of the macro description from the micro description is given and a connection between the two levels is provided. In my view, the connection is not an ontological identity relation: the aggregate state concept is not identical to the processes on the micro level. So the micro-reduction of this section does not identify the two different levels with each other, but as Looijen (1998), I think this is never possible. The micro-reduction in this section is an epistemological reduction, and if reduction means the formulation of an aggregate process concept describing the macro level in micro terms, then micro-reduction is successfully applied to the refined suspicion condition of the simulation case.

In this thesis I stressed the importance of paying attention to the reduction of the properties of a system. In the approach to the simulation case this resulted in a concept describing a process in the simulation case. The process however, can also be described by a law and then an approach of theory reduction is also possible. Maybe it is not possible to draw a strict line between concept reduction and theory reduction and a description of some processes might amount to the same thing as a law. But even if there is no strict border between concept reduction and theory reduction, it at least was possible to formulate the phenomena in process concepts.

4.3 Emergentistic approach

The characteristic patterns of suspicion and segregation in the simulation case are the macro phenomena. Out of the simple rules and interactions of the cellular automata, highly recognizable patterns emerge. The new patterns can be explained in higher order (social) terms. In this section

I will apply the emergence conditions I discussed in chapter 3 to the simulation case, to decide whether the macro patterns are emergent or not.

4.3.1 Simulation, complexity and non-aggregativity

To run the simulation is a way to discover the final outcome of a particular initial random distribution of the individuals over the cells. According to this first condition of emergence, the process is an emergent process if the only way to predict the outcome is by simulation. Predicting the outcome means to predict the exact positions of the individuals when they have reached a stable position, not just the prediction of the macro patterns. The exact positions differ among different runs of the simulation, whereas the global patterns remain constant over different runs. In contrast to the macro patterns, the exact positions cannot be predicted from a preference table alone. Whether simulation is the only way to determine the final outcome of the simulation is not certain. No algorithms or calculations to do this in fewer steps (given a preference table and the starting positions) are known in the literature, but a proof to the contrary, simulation is the only way to predict the outcome, is neither known. Although it does not give a lot of information, the ‘prediction-by-simulation-only’ condition does at least not contradict the possibility we can speak of emergence in this case.

The second possible criterion of emergence is complexity, in which emergence requires complexity. In chapter 3 I discerned four degrees of complexity in different kinds of systems: fixed, periodic, complex and chaotic systems. The simulation case always reaches a stable distribution of the individuals over the cells, nothing changes after that. In the first complexity category the relationships between the elements are always the same and the systems remain unchanged over time. So from a certain point in time the simulation case belongs to the first category of the classification, a fixed system. According to the complexity condition, fixed systems are not complex and thereby not emergent. Before this moment however, the system and the relations between its elements do change. This process is not an endless repetition of the same process, so it cannot be a periodic pattern. The process in the simulation neither is totally random, because it always leads to particular macro patterns. Then the conclusion is that the process before the simulation reaches a stable outcome is complex and thereby emergent. So the complexity condition of emergence classifies the process in the simulation case as an emergent process, the resulting macro pattern alone cannot be referred to as emergent.

Before I apply the third possible condition of emergence to the simulation case, I will first repeat Wimsatt’s four conditions of aggregativity.

- 1 Intersubstitutivity of parts
- 2 Qualitative similarity with a change in the number of parts
- 3 Stability under reaggregation of parts
- 4 Minimal interactions among parts

(Wimsatt, cited in Bechtel 2003)

The *first condition* of aggregativity is not totally satisfied. The gray individuals show identical behavior and could be intersubstituted without consequences and the same applies to the black individuals. But gray and black individuals cannot be intersubstituted on both sides without having effects on the final outcome of the simulation case.

The simulation case does answer Wimsatt's *second condition*, that is, within certain limits the total number of the cells and the individuals can be changed. In the first section of this chapter I described an early and a larger version of the simulation case. The two versions of the simulation had different numbers of cells and individuals, but they still showed the same results. The results were clearer in the larger model, but there was no qualitative difference. So the condition 'qualitative similarity with a change in the number of parts' is satisfied.

The *third condition* asks whether stability under reaggregation of parts is present. One of the rules of the simulation case requires a random distribution of the individuals over the cells at the start. Many, if not all, different initial distributions of the individuals lead to the same macro patterns in the end situation. So this condition of aggregativity is fulfilled.

The *last condition* of minimal interactions is not satisfied. In the simulation case individuals are influenced by other individuals, they can attract and reject each other. This suggests interaction among the parts, which is not a feature of aggregativity according to Wimsatt.

The conclusion Wimsatt would draw from this non-aggregativity condition is that the simulation case involves emergent phenomena. Failing one condition is already enough to speak of emergence, here two of the four conditions of aggregativity (condition one and four) are not satisfied. In chapter 3 I added a restriction to Wimsatt's four conditions, I also required the emergent phenomena to be processes to exclude static emergence. This demand is partly fulfilled, during a certain amount of time the simulation case obviously is a process and involves interactions between the parts.

4.3.2 Boundary conditions

A system is embodied when it has a body with capabilities for sensing and moving. Agents in a simulation cannot have real bodies, but in the simulation case of this thesis bodies and capabilities of the individuals are simulated. Individuals can move over the grid (one step or one jump) and they 'sense' the presence of other individuals. Situatedness of the individuals is also simulated, they act in an environment with empty cells and cells occupied by other individuals. In the analysis of the simulation case, two levels of organization were discussed, a micro and a macro level. The individuals on the micro level are not organized out of smaller parts, so it is not possible to zoom in on them. It is possible to zoom out and view the whole system placed in an environment.

In the simulation case, the influence of the environment on the system on which this chapter concentrated could be called boundary conditions. Boundary conditions are for example the borders of the grid, individuals cannot fall off the grid. These macro boundary conditions determine a great part of the final shapes of the macro patterns. In the segregation condition for example the borders of the grid determine the forms of the islands, the shapes of the islands follow the borders of the grid instead of being round. A 'half island' is a description on the macro level, but it is also possible to translate the macro boundary conditions into the micro level. Single individuals cannot move across the borders and near to a border less moves are allowed (five or even three instead of eight possibilities). The causal effects of the macro boundary conditions correspond on the micro and the macro level, this supports the view that there is a connection between the two levels in some way.

Because of the specific boundary conditions, particular patterns arise. By viewing the simulation case from a macro perspective (including the macro boundaries), regularity in the patterns can be recognized. These patterns would not be recognizable with access to only a part of the grid. Therefore the simulation case shows that the environment plays an important role in emergence. Macro processes are not only determined by the interactions between the corresponding micro parts, but also by environmental conditions.

4.3.3 Conclusions

The macro phenomena in the simulation could be emergent according to the first condition of emergence (prediction by simulation only), this condition did not contradict that possibility. The complexity condition classified the process of evolving to the final macro pattern as emergent. The conditions of non-aggregativity also classified some of the macro phenomena as emergent. None of the three conditions alone is highly convincing, but together they provide a strong argument that the simulation case involves emergent phenomena.

In the previous subsection I concluded with some notes on the importance of the environment in the notion of emergence. To my opinion, the analysis of the simulation case contributes to this view and shows that the environment plays an important role in the development of emergent properties.

5 Conclusion

In this final chapter the different parts of the thesis will come together. From a comparison between the reductionistic and emergentistic approach to the simulation case, I will draw conclusions about reduction and emergence in general. Then I will go back to the mind-body problem and as far as possible provide an answer to the main question of this thesis.

5.1 Reduction, emergence and the simulation

The micro-reduction of the simulation case in chapter 4 gave an explanation of how the rules and properties of the micro objects yielded the macro patterns in the simulation case. In the emergentistic approach I applied features of emergence to the simulation case and according to these conditions the simulation held emergent properties. So the simulation case in chapter 4 seems to be appropriate for an emergentistic and a reductionistic approach. However, not all reductionists and emergentists think that the two approaches are compatible in this way. Teller, for example, stated the following.

“An emergent property of a whole does not reduce to, in the sense of being explicitly definable in terms of, properties of the parts.” (Teller 1992, p140)

And Beckermann:

“The concept of micro-reduction is the exact complement of the concept of emergence” (Beckermann 1992, p 115)

According to Beckermann, for a property to be emergent it means that the property is not micro-reducible by definition. If he is right, it would mean that either the reduction I gave is not correct, or that the macro properties in the simulation case are not emergent. Looijen, Wimsatt and Bechtel share different opinions, they stated the following. According to Looijen:

“If there were no emergence, there would be no (micro-)reductions” (Looijen 1998, p79)

Wimsatt:

“A system property may be reducible to properties of its parts and their relations and still be spoken of as emergent” (Wimsatt 1986, p259)

And Bechtel:

“It is possible to be both a reductionist and an emergentist” (Bechtel 2003)

In my opinion, the simulation case contains emergent phenomena and a reductionistic approach is also possible, so I do not agree with Beckermann on his opinion that reduction and emergence exclude each other. I also think that the simulation case can be generalized and that reduction and emergence neither are excluding in other cases. To argue for reductionism and emergentism as supplementary instead of excluding approaches, I will first describe similarities between reductionism and emergentism to point out some common features. Then I will describe some differences between them and discuss why these differences do not mean the two approaches exclude each other.

5.1.1 Similarities between reduction and emergence

The views of emergentists and reductionists both involve two levels of description, a macro level and a micro level. The reductionists and emergentists I discussed claimed these two levels to be distinctive ontological levels. Reductionists refer to phenomena in these respective levels with ‘phenomena to be reduced’ and ‘reducing phenomena’, emergentists call the different phenomena ‘the emergent’ and ‘its parts’. The micro level of the simulation case is described by the behavior of the individuals on the grid: ‘the micro objects’ or ‘the parts’, the macro level is described by the characteristics of the macro patterns. Both approaches give an account of the relation between the two levels in the simulation case.

A common aspect that appeared in the discussion of both approaches to the simulation case was the emphasis on the processes in the simulation, instead of focusing on the stable end-situations. According to the complexity criterion of emergence, the simulation case was emergent up to the point a stable outcome was reached, after that the simulation case was no longer classified as emergent. In the section about reduction, I remarked that the formulation of an aggregate concept of the macro phenomena in the suspicion condition was not complete by just calling it ‘the sum of the distances’; this formulation only refers to a part of the characteristic macro phenomena in the simulation case. The formulation should take the process in the

simulation case into account by calling the macro phenomena ‘maximizing the sum of the distances’. Without this adaptation the reduction would not give an account to all the macro phenomena in the simulation case, it would only reduce a part of the macro phenomena. So in both approaches to the simulation case it became clear that a focus on the processes in the simulation was necessary.

Another similarity in the analyses of the simulation case deals with the role of micro properties in providing explanations. The aggregate concept formulated in the reductionistic approach was defined in micro terms. So a macro concept was tried to derive from properties of its micro objects. In his account of emergentism, Wimsatt approached macro phenomena by analyzing the way they were built up by their parts. So he also paid attention to the properties and interactions of the parts on the micro level. So a common feature in both approaches is the explanation of macro behavior by properties of and relations between the micro objects.

One more common feature between the reductionistic and the emergentistic accounts of micro-macro relations has to do with some comments I made on the condition of prediction by simulation only. According to some emergentists, simulating an emergent process is the only way to predict how the process will continue. My criticism on this feature of emergence had to do with the impossibility to check whether simulation really is the *only* way to make predictions or not. However, processes on a macro level can in theory always be predicted by simulation with complete knowledge of the micro level, so macro level phenomena can at least be derived from micro level phenomena by repeating all the steps of a process. In the simulation case of this thesis, no other way to predict the outcome was known than just executing all the steps. The reduction I gave was a description of all the steps that had to be taken in the simulation case. The induction steps taken in the reduction included all the steps that should be taken in a run of the simulation case. So as in the emergentistic approach, the reductionistic approach did not provide an account with an algorithm to describe the processes of the simulation case in fewer steps. If an outcome of an emergent process would be easier to predict (in fewer steps) than by simulation, then the reduction would probably also be shorter.

The final similarity between reductionism and emergentism involves the role of the environment. In both approaches the environment is needed to mark macro phenomena. Without an environment we would not be able to recognize macro systems as wholes, something is perceived as a unity or a whole because of its interactions with the environment. This is important for both reductionism and emergentism.

5.1.2 Differences between reduction and emergence

A first difference between reduction and emergence has to do with the distinction I made between process and static emergence and the choice to only take process emergence into account.

Because of this decision, according to the weak emergence used in this thesis, static macro phenomena are not emergent. The kinds of reduction used in the thesis however, do give accounts of how to reduce static macro phenomena. So reductionism does take static macro phenomena into account, like for example ‘being a circle’, whereas emergence does not. However, with another definition of emergence this difference would disappear. Even with the definitions used in this thesis, reductionism and emergentism are not mutually excluding approaches. It just makes reductionism broader than emergentism. The macro phenomena to which emergence is applicable are a subset of the macro phenomena appropriate for reductionism.

The two approaches to the simulation case are somewhat difficult to compare, because I had to answer different questions about the simulation case in both sections. In the section about reductions, I gave an answer to the question how the macro patterns can be reduced to the micro processes. I did this by giving a reduction, that is, a reductive derivation. In the emergentistic approach I gave an answer to the question whether the macro patterns were emergent patterns yes or no, I concluded a yes.

This difference has to do with different goals of the two approaches. Reductionists try to reduce the macro phenomena they observe into micro processes. Emergentists also see these macro phenomena and some of them (e.g. Wimsatt) even explain the macro phenomena by specific features of the micro parts and interactions. However, reductionists claim to prefer a micro level description over a description on the macro level, emergentists rather describe macro level phenomena in macro terms, they only investigate the way a macro level is built up to better understand it.

Closely related to the previous difference is the goal of reductionists to provide a connection between the micro and the macro level in their reductions by giving ontological identity relations. Emergentists on the contrary, investigate the relation between different levels of organization without the goal of identifying phenomena in the different levels with each other, they even think this is not possible. But Looijen, a reductionist, remarks that micro-reduction is not real concept reduction, because there cannot be an ontological identity relation between two phenomena on different ontological levels (Looijen 1998). Kuipers, also a reductionist, agrees with Looijen that in micro-reduction two epistemological levels are identified with each other, not the two different ontological levels (Kuipers 2001).

5.1.3 Conclusions

Emergentism and reductionism are obviously not the same, but to my opinion Beckermann's statement about their absolute incompatibility is too strong. The similarities I described show that both approaches are committed to multi level ontology and their explanations of macro properties by the micro properties have some features in common. The differences I described show that they have different methods with different ideas and goals, but these differences between reductionism and emergentism are not contradictions. Reductionism for example is broader applicable, it also applies to static phenomena whereas emergentism (as used in this thesis) does not. This difference does not point to a contradiction between the two theories. The preferences of emergentists and reductionists how to describe macro phenomena, in macro or in micro terms, neither indicates a contradiction.

So my conclusion is that reductionism and emergentism can go together, but I must stress the conditions under which I reached this conclusion. Emergentism and reductionism both appear in many different forms and not all forms are compatible with each other. Some emergentists claim that emergent phenomena are not derivable from complete information about the micro level and some kinds of reductionism do not involve multiple realizability or require ontological identification for a reduction. However, in my view I chose appropriate and reasonable accounts to analyze the simulation case and my conclusion at least applies to the forms of reduction and emergence I used in this thesis.

5.2 Dualism, identity theory and functionalism

Now it is time to turn back to my initial question concerning the mind-body problem. How does the discussion on reductionism and emergentism bear on this problem? In this section I will compare my view with the three theories about the mind and the body described in chapter 1: dualism, identity theory and functionalism.

I can be short about dualism, because in the first chapter already appeared that my view is not dualistic but physicalistic or materialistic. In my analysis, the processes of micro parts (neurons) explain the (mental) macro phenomena and nothing more than that exists. This strongly differs from a dualistic view.

The second theory of the mind is the identity theory, claiming that mental processes and neural processes are identical and can be reduced to each other. In my approach, macro

phenomena are also reduced to micro phenomena, but in a different way than in identity theory. According to identity theorists, mental and physical phenomena are on the same ontological level (Kim 1996), so the reduction between the two levels takes place on one and the same ontological level. They think that the reduction of mental phenomena to physical phenomena involves an ontological identity relation. In the reductionistic approach I used in this thesis, two ontological levels are involved in a micro-reduction. In the simulation case for example, the individuals on the micro level and the global patterns on the macro level both constituted their own ontology. In this thesis, a micro-reduction does not involve an ontological identity relation, macro levels are not ontologically identified with micro levels. Here, a micro-reduction consists of a description of the macro level in micro level terms. In this thesis I tried to make it convincing that the micro macro model also applies to the mind-body problem. The micro level is formed by the neurons and their interactions, the macro level is constituted by the more complex behavior of human beings as a whole. The conclusion is that my case contradicts strict identity and my approach is different from the identity theory of mind.

Functionalism is the last theory of the mind discussed. Its ideas of multiple realizability correspond with different runs in the simulation case, where different micro configurations lead to the same macro phenomena. Different distributions of the individuals over the cells lead to the returning patterns of for example suspicion and segregation. Functionalism and emergence both want to denote certain genuineness to these macro phenomena. Functionalists do this by describing the causal roles of a property, this makes (mental) macro phenomena what they are. Emergentists use different criteria to indicate emergent macro phenomena, for example the conditions of simulation by prediction only, complexity and forms of non-aggregativity. In the simulation case for example, the macro patterns do not have new causal powers, but still were denoted as emergent. So in contrast to functionalism, in my approach something can be emergent, without having specific causal roles or a specific function.

In the approach to the mind-body problem I suggest in this thesis, a human is seen as a system with different levels of organization and it makes use of reductionistic and emergentistic theories to analyse the relation between mind and body. This account differs from the three important theories in the mind-body problem: dualism, identity theory and functionalism. It is a physicalistic theory, it does acknowledge mental and physical phenomena as two different things and attention is paid to the way the mental is built up of the parts. It would be worthwhile to pursue this line of thinking in the context of philosophy of mind.

5.3 Main question

This last section provides a summary of the answers to the threefold main question, gathered in this thesis. The questions were:

(a) To what extent are higher order processes real and autonomous with respect to their underlying realizations, (b) do reduction and emergence describe the relation between the micro and the macro level in a different way, and (c) how does this bear on the mind-body problem?

The first question (question a) is to what extent are higher order processes real and autonomous with respect to their underlying realizations? Higher order processes are autonomous to a certain extent, because systems with higher order processes have properties and thereby causal powers its parts do not have. The higher order processes constitute a new ontological level, different from the ontological level of the underlying realizations. Nevertheless, higher order processes are not totally disconnected from their underlying realizations. Repeating a process with precisely the same conditions on the micro level will lead to the same processes on the macro level. Macro phenomena can also be described in micro terms, such a description should explain the way a macro system is built up.

The second part of the question, question b, asks in how far there is a difference between reductionism and emergentism to describe the relation between micro and macro levels. With the accounts of reductionism and emergentism used in this thesis the answer is the following. The goals of the two theories are different. Reductionists prefer micro level descriptions to macro level descriptions and emergentists describe the respective levels in their own terms. But reductionists and emergentists both agree that higher order processes constitute a different ontology compared to the underlying realizations. Both approaches also explain macro phenomena with properties of the micro objects in collaboration with the body and the environment.

The last part of the question, part c, is about how all this bears on the mind-body problem. The mind can be seen as an emergent phenomenon, it emerges out of neural micro processes and constitutes a new ontological level. Although it is impossible to ontologically identify the mind to the body, to my opinion it is possible to give a reduction of the mind to the body, at least by deriving the mental from the physical by simulating all the physical micro processes. Reductionism and emergentism supplement each other in analyses of systems with

different levels of organization. This joint approach, different from the identity theory and functionalism, can also be applied to the mind-body problem. To gain more insight about the mental and the physical and their relation, phenomena on all levels should be investigated and the environment should also be taken into account.

Bibliography

- Bechtel, W., Mandik, P., Mundale J. and Stufflebeam, R.S. *Philosophy and the Neurosciences*, Boston, Blackwell Publishers, 2001.
- Bechtel, W. Philosophy engages cognitive neuroscience (lecture 1 t/m 6), *retrieved*: <<http://mechanism.ucds.edu/~bill/research/mercier>>, 2003.
- Beckerman, A., Flohr, H. and Kim, J. *Emergence or Reduction? Essays on the prospects of Nonreductive Physicalism*, Berlin, De Gruyter, 1992.
- Bedau, M.A. Weak emergence, *Philosophical Perspectives*, 11, 375-399, 1997.
- Bedau, M.A. Downward causation and the autonomy of weak emergence, *Principia*, 6, 5-50, 2002.
- Bedau, M.A. Artificial life, *to appear: Blackwell Guide to the Philosophy of Computing and Information*, ed. Luciano Floridi, 197-211, Oxford, Blackwell Publishing.
- Bickle, J. *Psychoneural Reduction The New Wave*, Cambridge, MIT Press, 1998.
- Clark, A. *Mindware - an introduction to the philosophy of cognitive science*, Oxford University Press, 2001.
- Dennett, D.C. Real patterns, *The journal of philosophy*, 99 (1), 27-51, 1991.
- Dennett, D.C. *Freedom evolves*, New York, Penguin Books, 2003.
- Hegselmann, R. and Flache, A. Understanding complex social dynamics: a plea for cellular automata based modelling, *Journal of Artificial Societies and Social Simulation*, 1(3), *retrieved*: <<http://www.soc.surrey.ac.uk/JASSS/1/3/1.html>> ,1998.
- Huisman, L. New Wave Reduction and its Relation to Reality An analysis of Bickle's solution to the mind-body problem, *Research paper*, University of Groningen, 2005.
- Kim, J. *Philosophy of mind*, Boulder, Westview Press, 1996.
- Kim, J. *Mind in a physical world*, Cambridge, MA: MIT Press, 1998.
- Kim, J. The layered model: metaphysical considerations, *Philosophical Explorations*, V(I), 2-20, 2002.
- Keijzer, F.A. *Representation and behavior*, Cambridge, MIT Press, 2001.
- Kuipers, T.A.F. *Structures in science. Heuristic patterns based on cognitive structures*, An advanced textbook in neo-classical philosophy of science, *Synthese Library*, 301, Dordrecht, Kluwer Academic Publishers, 2001.
- Looijen, R. Holism and reductionism in biology and ecology, *PhD Thesis*, University of

- Groningen, 1998.
- Marraffa, M. and Nani, M. (general editors) Concepts of intertheoretic reduction, *retrieved*: <http://host.uniroma3.it/progetti/kant/field/cir.htm>, page last updated in 2002.
- McLaughlin, B. P. The rise and fall of British emergentism, *Emergence or Reduction*, eds. Beckermann, A., Flohr, H., Kim, J., 49-93, Berlin, De Gruyter, 1992.
- Nagel, E. *The structure of science*, New York: Harcourt, Brace & World, 1961.
- Pfeifer, R. and Scheier, C. *Understanding Intelligence*, Cambridge, MA: MIT Press, 1999.
- Sakoda, J. M. The checkerboard model of social interaction, *Journal of Mathematical Sociology*, 1, 119-132, 1971.
- Salen, K. and Zimmerman, E. *Rules of Play*, Cambridge, MA: MIT Press, 2004.
- Smart, J.J.C. The identity theory of mind, *Stanford Encyclopedia of Philosophy*, *retrieved*: <http://plato.stanford.edu/archives/fall2004/entries/mind-identity/>, 2004.
- Stephan, A. Emergence – A systematic view on its historical facets, *Emergence or Reduction*, eds. Beckermann, A., Flohr, H., Kim, J., 25-48, Berlin, De Gruyter, 1992.
- Teller, P. A contemporary look at emergence, *Emergence or Reduction*, eds. Beckermann, A., Flohr, H., Kim, J., 139-153, Berlin, De Gruyter, 1992.
- Wimsatt, W.C. Forms of aggregativity, *Human Nature and Natural Knowledge*, eds. Donagan, A., Perovich, N. and Wedin, M., 259-291, Dordrecht, Reidel, 1986.
- Wimsatt, W.C. Aggregativity: reductive heuristics for finding emergence, *Philosophy of Science*, 64, 372-384, 1997.