

Paradox in AI – AI 2.0: The Way to Machine Consciousness

Peter Palensky¹, Dietmar Bruckner², Anna Tmej², and Tobias Deutsch²

¹ University of Pretoria, South Africa
palensky@ieee.org

² Vienna University of Technology, Austria
{bruckner,tmej,deutsch}@ict.tuwien.ac.at

Abstract. Artificial Intelligence, the big promise of the last millennium, has apparently made its way into our daily lives. Cell phones with speech control, evolutionary computing in data mining or power grids, optimized via neural network, show its applicability in industrial environments. The original expectation of true intelligence and thinking machines lies still ahead of us. Researchers are, however, optimistic as never before. This paper tries to compare the views, challenges and approaches of several disciplines: engineering, psychology, neuroscience, philosophy. It gives a short introduction to Psychoanalysis, discusses the term consciousness, social implications of intelligent machines, related theories, and expectations and shall serve as a starting point for first attempts of combining these diverse thoughts.

Keywords: Machine consciousness, artificial intelligence, psychoanalysis.

1 Introduction

Embedded computer systems have seen their computing power increase dramatically, while at the same time miniaturization and wireless networking allow embedded systems to be installed and powered with a minimum of support infrastructure. The result has been a vision of “ubiquitous computing”, where computing capabilities are always available, extremely flexible and support people in their daily lives.

The common view in related communities is that computers will not only become cheaper, smaller and more powerful, but they will disappear and hide or become integrated in normal and everyday objects [21], [22]. Technology will become invisible and embedded into our surroundings. Smart objects communicate, cooperate and virtually amalgamate without explicit user interaction or commands; they form consortia for offering or even fulfilling tasks for a user. They are capable of not only sensing values, but deriving context information about the reasons, intentions, desires and beliefs of the user. This information may be shared over networks – one of which is the world-wide available Internet – and used to compare and classify activities, find connections to other people

and/or devices, look up semantic databases and much more. The uninterrupted information flow makes the world a global village and allows the user to access his explicitly or implicitly posed queries anywhere anytime.

The vision of machine consciousness – which includes the availability of enough computation resources and devices for fulfilling the tasks – poses many requirements across the whole field of information and communication technology and allied fields [23]. For example, the development in ambient intelligence with respect to conscious environments requires development in the areas of sensors, actuators, power supplies, communications technology, data encryption and protection, privacy protection, data mining, artificial intelligence, probabilistic pattern recognition, chip design and many others not stated here. Each research group and even researcher has its own view on what machine consciousness will be or will approximate. The same is already true for the many necessary technologies required for machine consciousness. The research on ambient intelligence for instance can be divided up [24] into three basic research areas – or existing research projects can be grouped into projects investigating three basic methods – which reflect fundamentally different approaches in establishing ubiquitous computing environments: Augmented Reality (see for example [25]), Intelligent Environments (e.g. [26], [27], [28]), and Distributed Mobile Systems (see [29], [30]).

The border between smart devices and devices with machine consciousness may be something very controversial and from the philosophic point of view of massive importance and implications. This will be true for ambient intelligence applications as well as any other, e.g. industrial, conscious application. However, for the user of such a device this distinction is meaningless: A user expects the fulfillment of a particular task, the device just needs to be intelligent enough to do so. An argument for calming down the apprehension of machines that will control us is presented below with the model of nested feedback loops, each loop representing a bit more consciousness than the lower ones. In this context the designer of a machine can control the level of consciousness he wants from his device.

These considerations together give a rough definition of machine consciousness as we see it: Machines equipped with decision units that allow them to think in a way humans do. The key issue is the “thinking in a way humans do” in opposition to “acting looking like humans” of many well-known projects, which is nothing but mimicry. There are, e.g., robots capable of performing facial expression like smiling. However, this does not imply that the robot “is amused” as a human would be who smiles, or that the robot has any other human-like intentions to smile. But this would be the necessary requirement for attributing this robot consciousness.

2 What Consciousness?

Consciousness¹ helps us humans to put ourselves into the set of parameters and actors when decisions are made. A concept of self, its desires and plans is one attribute that makes someone appear intelligent.

¹ The interested reader might be referred to David Chalmers’ and David Bourget’s web repository <http://consc.net/online>, which lists more than five thousand papers on consciousness.

Most of all, consciousness is a subjective quality. Person A can feel its own consciousness, be aware of it, test it and have a concept of itself. Person A can tell Person B about this extraordinary experience but Person B can never be sure. Only from “inside”, consciousness can be experienced and verified. The link of this qualia to the physical process, the physical machine, is unfortunately not clear [35].

The outer shell of a conscious being might exhibit very distinctive behavioral patterns. These patterns might in turn be checked against a “turing test on consciousness”. There are, however, numerous arguments about machines, potentially passing these tests but still being “zombies”, i.e. lacking consciousness [36].

Another distinction is where we assume consciousness. Generally accepted as a human phenomenon (maybe also in “higher” mammals), its projection to or implementation in machines [37] opens up two problems:

1) We could be fooled by anthropomorphic aesthetics. Humans actively seek emotions in faces, even if they are inanimate rubber masks and we should be aware of this pitfall.

2) If we created this machine, we have total power over its hardware (and software, if we use contemporary terms) and can copy it, switch it off and on without damage, modify it, monitor it without interference, etc.

It is the second point that would make these beings massively different to humans. We have no power over our own hardware. We are mortal and do not fully understand how our body works. Out of this situation we can derive two alternative outcomes:

a) A machine that can be switched off and on cannot host consciousness because it would be far too primitive. Once we have machine consciousness we will realize that the hardware has been gradually taken out of our hands (e.g. manufactured by nano-bots and evolutionarily modified beyond our knowledge) and therefore as intangible as our own brains.

or

b) It works and constitutes one half of what many people dream of: a potential “storage” or platform for our own consciousness, ultimately leading to immortality (the other half, however, is still missing: understanding of our own hardware a.k.a. “wetware”, to download its content to the new platform). Unfortunately, a subjective engineering estimate would be that the chances that such a machine – if it develops something like consciousness – is compatible to our “software” are virtually zero.

So we might reach some point where we have remarkable and potentially conscious artifacts, but their comparison to our own body and mind as a unity is very questionable, although most technical applications would be happy with the mind of a sophisticated “zombie” [42].

3 Consciousness Support

To provide a machine with machine consciousness one must, in order to stay scientifically consistent, review already existing research on human consciousness.

In this chapter we will give an overview about how various sciences approach consciousness. For technical purposes a holistic, functional model without contradictions is desired. We will choose an existing method as a template and go into more depth into that theory.

3.1 Philosophy

The questions of consciousness that philosophy traditionally dealt with can be gathered in three crude rubrics as the What, How and Why questions: What is consciousness, its principal features? How does consciousness come to exist, and finally, Why does consciousness exist? David Chalmers [7] summarizes more modern philosophical approaches when he differentiates between the so-called “hard problem” and the “easy problem” of consciousness. The latter concerns objective mechanisms of the cognitive systems: the discrimination of and reaction to sensory stimuli, the integration of information from different sources and the use of this information to control behavior and the verbalization of internal states, while it is the “hard problem” that actually deals with the “mystery” of consciousness [7], p. 62: the question of how physical processes in the brain give rise to subjective experience. This involves the inner aspect of thought and perception: the way things feel for the subject. This part of consciousness is also called “phenomenal consciousness” and “qualia” [8], while Chalmers’ “easy problem” is also called “access-consciousness”. Daniel Dennett [9], on the other hand, denies that there is a “hard problem”, asserting that the totality of consciousness can be understood in terms of impact through behavior. He coined the term “heterophenomenology” to describe an explicitly third-person scientific approach to human consciousness.

Dennett’s new term is closely linked to his cognitive model of consciousness, the Multiple Drafts Model (MDM) [9]. According to this model, there are a variety of sensory inputs from a given event and also a variety of interpretations of these inputs. The sensory inputs arrive in the brain and are interpreted at different times, so a given event can give rise to a sequence of discrimination, constituting the equivalent of multiple drafts of a story. As soon as each discrimination is accomplished, it becomes available for eliciting a behavior. Like a number of other theories, the Multiple Drafts Model understands conscious experience as taking time to occur. The distinction is that Dennett’s theory denies any clear and unambiguous boundary separating conscious experiences from all other processing. According to Dennett, consciousness is to be found in the actions and flows of information from place to place, rather than some singular view containing our experience. The conscious self is taken to exist as an abstraction visible at the level of the intentional stance, akin to a body of mass having a center of gravity. Similarly, Dennett refers to the self as the center of narrative gravity, a story we tell ourselves about our experiences. Consciousness exists, but not independently of behavior and behavioral disposition, which can be studied through heterophenomenology.

3.2 Psychology

Before the advent of cognitive psychology, psychology failed to satisfactorily study consciousness. Introspectionism, first used by Wilhelm Wundt as a way to dissect the mind into its basic elements, dealt with consciousness by self-observation of conscious inner thoughts, desires and sensations. The relation of consciousness to the brain remained very much a mystery. This experimental method was criticized by upcoming behaviorists as being unreliable; scientific psychology should only deal with operationalizable, objectifiable and measurable contents. Behaviorism thus studied the mind from a black-box-point-of-view, stating that the mind could only be fully understood once the inputs and the outputs were well defined, without even hoping to fully understand the underlying structure, mechanisms, and dynamics of the mind.

Since the 1960s, cognitive psychology has begun to examine the relationship between consciousness and the brain or nervous system. The question cognitive psychology examines is the mutual interaction between consciousness and brain states or neural processes. However, despite the renewed emphasis on explaining cognitive capacities such as memory, perception and language comprehension with an emphasis on information processing and the modeling of internal mental processes, consciousness remained a largely neglected topic until the 1980s and 90s.

A major example of a modern cognitive approach to consciousness research is the global workspace theory of Bernard Baars [1]. It offers a largely functional model of consciousness which deals most directly with the access notion of consciousness and has much in common with the multiple drafts model mentioned above. The main idea of global workspace theories is that consciousness is a limited resource capacity or module that enables information to be “broadcast” widely throughout the system and allows for more flexible sophisticated processing. It is thus closely allied with many models in cognitive psychology concerned with attention and working memory.

Just as the philosophical approaches mentioned above, however, psychological attempts at describing and investigating consciousness lack a thorough and exact analysis of the structure, mechanisms and dynamics of consciousness and mental processes.

3.3 Evolutionary Biology

Evolutionary Biology has mainly investigated the question of the causes for consciousness. From this point of view, consciousness is viewed as an adaption because it is a trait that increases fitness.

3.4 Physics

Modern physical theories of consciousness can be divided into three types: theories to explain behavior and access consciousness, theories to explain phenomenal consciousness and theories to explain the quantum mechanical (QM) Quantum

mind [43], [44]. These latter theories are based on the premise that quantum mechanics is necessary to fully understand the mind and brain to explain consciousness. The quantum mind hypothesis proposes that classical mechanics cannot fully explain consciousness and suggests that quantum mechanical phenomena such as quantum entanglement and superposition may play an important part in the brain's function and could form the basis of an explanation of consciousness.

3.5 Cognitive Neuroscience

This scientific branch is the most modern approach to consciousness research, primarily concerned with the scientific study of biological substrates underlying cognition, with a specific focus on the neural substrates of mental processes and their behavioral manifestations. Amongst other things, it investigates the question of how mental processes uniquely associated with consciousness can be identified. It is based on psychological statistical studies and case studies of consciousness states and the deficits caused by lesions, stroke, injury, or surgery that disrupt the normal functioning of human senses and cognition. Cognitive neuroscience is a branch of both psychology and neuroscience, unifying and overlapping with several sub-disciplines such as cognitive psychology, psychobiology and neurobiology. One major question that cognitive neuroscience deals with is the so-called “mind-body-problem” [17], the question how brain and mind or brain and consciousness relate to each other. Many cognitive scientists today hold the view that the mind is an emergent property of the brain: mind and brain equally exist, however they exist at different levels of complexity.

Antonio Damasio [2] differentiates between “core consciousness” and “extended consciousness”. While core consciousness describes a hypothesized level of awareness facilitated by neural structures of most animals, which allows them to be aware of and react to their environment, extended consciousness is a much more complex form of consciousness, allowing for a sense of identity and personality and meta-consciousness, and linking past, present and future. Higher forms of extended consciousness only exist with humans and depend on (working) memory, thinking and language. Just as all other scientific approaches mentioned above, cognitive neuroscience, too, does not provide us with a detailed and concrete model of the mental processes involved in the structure and mechanisms of consciousness.

3.6 Psychoanalysis

Psychoanalysis opened up the research field of consciousness to the unconscious properties of the mind. As described in more detail below, Freud developed his metapsychological ideas from a theory of three mental processes – conscious, preconscious and unconscious – to a theory of three agencies – the Ego, the Id and the Super-Ego – functioning on the basis of these processes. Although Freud developed his theory as a contrast to the psychology of consciousness exclusively current at his time [10], naturally, the mind that Freud set out to analyze and describe with great precision includes and determines consciousness: although

many thoughts or other psychic/mental contents may never reach consciousness, they will always exert influence on consciousness and behavior. According to Freud, it is therefore both legitimate and necessary to include those properties that lie behind consciousness within our conception of the mind [17].

Freud and other psychoanalysts after him saw and see consciousness largely as a means to perceive outer and especially also inner events and inner-psychic qualities [31], as a property of the mind as opposed to the mind itself [17]. The mind, the mental apparatus, including conscious and unconscious properties was described by Freud in all its functions and dynamics. In the end, psychoanalysis emerges as the only science dealing with human consciousness on a detailed enough level to be used for an implementation of machine consciousness.

4 Psychoanalysis, the Template

For the scope of this work psychoanalysis was chosen as the template theory of human consciousness. The main reasons therefore are the functional approach of modelling human thinking, which very well fits a computer engineering approach, and the somewhat technical or natural scientific approach of Freud to the topic which resulted in texts and arguments that can be followed and agreed on by natural scientists. This chapter gives an introduction to people not from the field.

4.1 Basics

Psychoanalysis was founded by Sigmund Freud (1856-1939), originally a neurologist and neuroanatomist [6], who in his study of hysteria developed his first ideas about unconscious affects, psychic energy and the cathartic effect of verbal expression. In the course of his life, Freud developed these and other concepts further, abolished some, modified others, while inspiring many other scientists to join him in his quest. Some of these came to disagree with Freud in time and went on to pursue their own strands of theory.

Freud himself [12] described psychoanalysis as follows: “Psycho-analysis is the name (1) of a procedure for the investigation of mental processes which are almost inaccessible in any other way, (2) of a method (based upon that investigation) for the treatment of neurotic disorders and (3) of a collection of psychological information obtained along those lines, which is gradually being accumulated into a new scientific discipline” (p. 235). He goes on to differentiate: “The assumption that there are unconscious mental processes, the recognition of the theory of resistance and repression, the appreciation of the importance of sexuality and the Oedipus complex – these constitute the principal subject-matter of psycho-analysis and the foundations of its theory” (p. 250).

After Freud died in 1939, psychoanalytic theory and practice continued to be further developed by scientists and practitioners as Heinz Hartmann, Anna Freud (Sigmund Freud’s youngest daughter), and Melanie Klein, to name just very few. However, already before Freud’s death, there had been disagreements about the meaning of certain concepts or ways of treatment. Since then, controversy has continued to be characteristic of psychoanalytic theory; distinct schools

of thought focusing on different topics, e.g. Ego-psychology or object relations have emerged. Even today, more than 100 years after its foundation, psychoanalysis is a living science in the sense that the different concepts of the different schools are still being further developed and discussed. While psychoanalysis is for a great part concerned with psychotherapeutic methods, psychopathologies and individual developments leading thereto, one great part also consists of metapsychology, or conceptualizing the psychic apparatus and the way it (mal-) functions. As this part of the psychoanalytic theory shall be our main focus, this introduction will simply exclude the other, albeit certainly very important and also characteristic aspects of psychoanalysis.

Additionally, although, as described above, there are many different strands of theory in psychoanalysis, each with different main focuses and different understandings of certain concepts, in this introduction, we will still concentrate on Freud's original conception of psychoanalytic theory. This is due to the very fundamental and basal nature of Freud's original ideas which to this day have not been abolished, and continue to be valid also in modern psychoanalytic theory, although different schools may focus on different concepts from different periods in Freud's life and development of the theory. Details about the technical conceptualization and partly implementation thereof can be found in [40]. The authors there in a first attempt concentrated on employing the original psychoanalytic theory, before venturing further to include other, more modern psychoanalytic concepts, be it the case that these prove as usable and expedient as Freud's conception of the psychic apparatus.

4.2 The Mental Apparatus (1): The Topographical Model

Freud conceptualized the mind as being divided up into different parts, each the home of specific psychological functions. His first model, the topographical model [14], divides the mind into the unconscious, the preconscious and the conscious system. Characteristic of these different parts are two different principles of mental functioning [11] – the so-called primary and secondary processes. While secondary process thinking – typical of conscious processes – is rational and follows the ordinary laws of logic, time and space, primary process thinking – typically unconscious – is characteristic of dreaming, fantasy, and infantile life in which the laws of time and space and the distinction between opposites do not apply. Some psychological processes are however only unconscious in the descriptive sense, meaning that the individual is not aware of them, but they are easily brought to mind (preconscious). With dynamically unconscious processes, on the other hand, it is not by simple means of effort or change of attention that they can be rendered conscious. These to the conscious system unacceptable psychic contents are subject to repression and operate under the sway of primary processes.

4.3 Drive Theory

Freud saw the internal world as dominated by man's struggle with his instincts or drives. In his initial formulation of instincts, Freud [15] distinguished between

self-preservative (e.g. hunger) and sexual drives (libido). Later, he stressed the difference between sexual drives and aggressive or destructive drives (Eros vs. Thanatos) [16]. Classically, instinctual wishes have a source, an aim, and an object. Usually, the source of the drive is infantile and lies in the body, possibly in an erogenous zone. Over time, after several similar (real or imagined) satisfactions of instinctual wishes, source, aim and object begin to mesh together into a complex interactional fantasy, part of which is represented in the system unconscious.

4.4 The Mental Apparatus (2): The Structural Model

In the structural model, Freud [12] proposed three parts or structural components of the human mind: Id, Ego and Super-Ego.

Id

The Id is the first psychic structure of the mental apparatus, out of which, in the course of infantile and childhood development, the Ego and the Super-Ego evolve. It contains the basic inborn drives and sexual and aggressive impulses, or their representatives. As such, it is an inexhaustible source of psychic energy for the psychic apparatus: The Id's wishes strive for immediate satisfaction (pleasure principle) and therefore drive the functions of the Ego to act. Its contents are unconscious and function following the primary process.

Ego

In the course of infantile development, the perceptive and executive parts of the Id, responsible for drive satisfaction by perceiving the baby's environment and ways to gain satisfaction, start to form a new part of the mental apparatus: the Ego. The mature Ego's tasks, however, exceed perception and execution: the Ego has to control the primitive impulses of the Id and to adapt these to outer reality (reality principle) as well as to mollify the requirements of the Super-Ego. For these purposes, the Ego makes use of so-called defense mechanisms such as repression to keep unacceptable impulses within the Id and therefore evade conflict with either outer reality or Super-Ego-requirements. The contents of the Ego are partly conscious, partly unconscious. A however certainly not complete list of Ego functions should not omit the following [5]: consciousness; sensory perception; perception and expression of psychic agitation; thinking; controlling motor functions; memory; speech; defense mechanisms and defense in general; fighting, controlling and binding drive energy; integrating and harmonizing and reality check.

Super-Ego

The Super-Ego comprises the conscience and ideals, thus allocating (moral) rules and prohibitions which are derived through internalization of parental or other authority figures, and cultural influences from childhood onwards. The Super-Ego resembles the Ego in that some of its elements are easily accessible for consciousness, while others are not. Super-Ego ideation also comprises rational and mature presentations up to very primitive and infantile ones. The task of

the Super-Ego is to impact on the actions of the Ego, especially to support it in its defensive actions against the drives with its own moral rules. However, the relationship between the Ego and the Super-Ego will not always be this harmonic: in other cases, e.g. if the difference between instinctual or other repressed wishes from the Id and moral rules from the Super-Ego becomes too great, the Super-Ego produces feelings of guilt or a need for punishment inside the Ego.

4.5 Psychoanalysis and Neuroscience

In recent years, a new scientific research strand has developed aiming at supporting and reassigning psychoanalytic concepts by neuroscientific findings: neuropsychology. Solms [18] provides a summary of the neuroscientific support for some basic psychoanalytic concepts, one of which is the notion that most mental processes occur unconsciously. Different memory systems have been identified, some of which are unconscious, which mediate emotional learning. The hippocampuses, which lay down memories that are consciously accessible, are not involved in such processes. Therefore, no conscious memories are available and current events can “only” trigger remembrances of emotionally important memories. This causes conscious feelings, while the memory of the past event remains unconscious. In other words, one consciously experiences feelings but without conscious access to the event that triggers these feelings.

The major brain structures for forming conscious memories do not function within the first two years of life. Developmental neurobiologists largely agree that early experiences, especially between infant and mother, fundamentally shape our future personality and mental health. Yet none of these experiences can be consciously remembered. It becomes increasingly clear, that a great deal of our mental activity is unconsciously motivated.

Freud’s ideas regarding dreams – being instigated by the drives and expressing unconscious wishes – were at first discredited when rapid-eye-movement sleep and its strong correlations with dreaming were discovered. REM sleep occurred automatically and was driven by acetylcholine produced in a “mindless” part of the brain stem, which has nothing to do with emotion or motivation. Dreams now were regarded as meaningless, simple stories concocted by the brain under the influence of random activity caused by the brainstem. Yet recent work has revealed that dreaming and REM sleep are dissociable while dreaming seems to be generated by a network of structures centered in the fore-brains instinctual-motivational circuits. These more recent views are strongly reminiscent of Freud’s dream theory.

These developments and the advent of methods and technologies (e.g.: neuroimaging) unimaginable 100 years ago make it possible today to correlate the psychoanalytic concepts, derived from observation and interpretation of subjective experiences, with the observations and interpretations of objective aspects of the mind studied by the classical neurosciences, thus rendering psychoanalysis and psychoanalytic theories objectifiable.

5 Road-Map

The capabilities of the desired machine consciousness introduced above represent a development stage in the future, which requires several disruptive innovations. We cannot say today how long it will take to reach them. There are many other authors predicting human-like intelligence within some decades, or only in 100 years, or even never. The authors here believe we are much closer! We see many promising approaches to problems of AI, their only drawback lies in their position within the framework of the application – which, in our opinion, shall be given by psychoanalysis [40]. However, psychoanalysis is not concerned with all the necessary functions, it is just and primarily concerned with psychic functioning itself.

Psychoanalysis requires so-called memory traces. A memory trace is the mental representation of something, be it a person, an object, event or whatever. Psychoanalysis is not concerned with perception (which generates memory traces), or with direct motor control. For technical systems these are however key requirements, needless to say these are the fields where robotics, AI (e.g. Software Agents) and others put most of their efforts in, but we need to search for other templates than psychoanalysis. Candidates here are developmental psychology and neuroscience, to mention just two. It becomes more and more clear that a system showing something like machine consciousness needs to be designed following a model, of which the mentioned disciplines have different viewpoints. Therefore, it needs to combine them in a holistic manner.

Psychoanalysis together with the other mentioned fields span a huge search space for engineers. Therefore it is not possible (and useful) to start trying to implement all of it. It is necessary to define a road map with milestones and priorities. A good hint is given by Rodney Brooks in his recent Spectrum article [34], where he introduced four capabilities of children, which are vital behavioristic observable capabilities of human-like intelligent machines. Unfortunately, these have nothing to do with psychic functioning; they are just observations of behavior. There may however be one connection in the last point – the “theory of mind” which in psychoanalysis is called “mentalization” or “reflective functioning” [39], which is definitely a crucial capacity within AI (being able to guess what another person is thinking, planning; understanding why a person does whatever it is they are doing etc.).

5.1 Mental Functions

Before mentioning and commenting Brooks’ list, it is necessary to explain some of the very basic concepts of psychoanalysis that need to be addressed by the first machines with comparable behavior. These are the primary and secondary process, terms that refer to the quality of thoughts – or mental content in terms of uncontrolled, impulsive actions or considered ones. The functions necessary for this according to [19] are:

perception: The mind knows two kinds of perception, internal and external. Internal perception results from an observational perspective on the mind. Both together allow for mental content of the kind: “I am experiencing this”.

memory: The mind recognizes previous mental experiences of the mentioned kind. It is also able to derive cause-and-effect sequences thereof. These two capabilities together form the immature ego, which only depends on drives and environment.

emotions: The very basic function of emotion is to rate something in terms of good or bad in a biological sense. More biologically successful actions are felt satisfying. Emotions are used to rate the above mentioned perceptions and memory content. In this way, quantitative events acquire quality.

feelings: Feelings are the base for consciousness. Their function is to evaluate the other mechanisms. The kind of mental content that results is “I feel like this about that”. Feelings are also stored together with their generating mental experiences. They can be used to create motivation by trying to repeat previous experiences of satisfaction. Motivation searches through past experiences and matches them with the present to come to decisions about what is to be done. The decisions lead to actions (motor output). These all together form the primary process.

inhibition: Experience about actions show that some of them are only satisfying in the short-term, but not in the long-term. Therefore it is necessary to be able to tolerate temporary displeasure via inhibiting the immediate action plan. This capacity permits thinking. Thinking is inhibited (imaginary) action, which permits evaluation of potential (imaginary) output, notwithstanding current (actual) feelings. This function is called secondary process. It replaces immediate actions with considered ones.

The ability of the secondary process develops very early in a child at the age of around two. However, the finesse and long-term anticipation capabilities in imaginary actions develops during the whole life.

The above description lets us conclude that the mind is organized as a system of nested feedback loops. If we leave memory aside, the first level consists of perception and evaluation. Let’s call the pure perception mental content level 1². The evaluation together with the perception, towards it is targeted, are again stored. This mental content is another piece of information; let’s call it mental content level 2. Content of level 2 is rated with feelings. These together form mental content level 3. The concept of evaluating mental content with feelings from level 3 upwards is the same for the all upper layers. It is not clear at this point if there are intermediate levels, but one of the next levels is related to language. With this, some mental content reaches a level of attention which is important enough to give it a name or a phrase – and therefore being able to share it with others. Again, in the upper levels, reasoning and feeling about mental content related with language – dialogs, names, etc. – form the next levels.

It is important that in this concept the lower levels of some mental content need to be “created” before upper levels can even be thought of. This may be

² Here we need to mention that “perception” itself requires lots of computation before generating the above mentioned memory traces. So, in a holistic model, we would start with level 1 on the sensor level.

one reason why it is so hard to bring findings or knowledge of different context together.

5.2 Example

As an example let us think about the following situation: Somebody (being it a child, an ancient hunter, whoever) is walking in nature and sees a steep hillside, an edge. Mental content level 1 would be “I see an edge”. For some reason (maybe searching for food, playing, etc.) this edge generates the interest of the person, letting it think (conscious or not) “I am interested in that edge”. Therefore, the person moves itself to that place. This involves basic action planning. However, when the person reaches the edge, the region behind gets visible. Let us further assume the person sees something he has searched for. Then mental content of the kind “I like this place – and also the landmark (the edge) which indicates the place” could arise, because walking there results in reaching something desired. If this desired thing is interesting or important enough, the person could think “I would like to tell my spouse of this” and “I call this region (e.g.) the XYZ place and the landmark the RST edge”.

In this way, perception is enhanced with evaluation level by level. The more interest (positive or negative) it creates, the more chances for further processing are implied. Things which cannot be perceived (because of a lacking template) or which generate neutral evaluation are considered unimportant.

In the following the four capabilities mentioned by Brooks are described (the first statements of the following sections marked with quotes). We try to formulate necessary psychic functions related to the introduced levels of mental content that enable the observation of the desired behavior. Additionally, other necessary concepts for rudimentary machine consciousness beyond Brooks’ ones are introduced.

5.3 The Object-Recognition Capabilities of a 2-Year-Old Child

“A 2-year-old can observe a variety of objects of some type – different kinds of shoes, say – and successfully categorize them as shoes, even if he or she has never seen soccer cleats or suede oxfords. Today’s best computer vision systems still make mistakes – both false positives and false negatives – that no child makes.”

Children recognize shoes, not their size, weight, etc. They recognize shoes because of their roughly similar shape and their functionality. That means: children know what shoes in general look like (foot-shaped plus a place where the foot can be inserted) and what they are for (that is: for putting them on and walking around). The rest is variable: exact shape, color, price; and does not interfere with basic shoe-recognition. Therefore, also a giant, two-meter shoe placed on the top of the roof of a giant shoe-shop will be recognized as a shoe, although it technically is, of course, not a shoe (but: it is foot-shaped, even though giant, and you could theoretically put your foot into it and therefore also walk around in it); however the child is naive enough not to know about other things that would hinder this shoe from really functioning in a shoe-like manner.

In visual object recognition requirements this would translate into many problems, an algorithm would need to solve: Finding out the geometry of the picture. How large is the area covered, how large are the objects. There needs also to be a model of the object in 3D. The visual classifier has to be based not only on features representing the spatial change in e.g. luminosity of pattern, but needs to take contours into consideration. A promising approach in this respect could be a combination of the boosted cascade of feature detection [20] relying on a set of features like shape, color, etc., not only pattern, as described e.g. by [41].

The very important aspect however is deriving the meaning of things for the machine. What can the machine do with that object? Can it be evaluated good or bad (or some more categories) in terms of the application of the machine?

5.4 The Language Capabilities of a 4–Year–Old Child

“By age 4, children can engage in a dialog using complete clauses and can handle irregularities, idiomatic expressions, a vast array of accents, noisy environments, incomplete utterances, and interjections, and they can even correct nonnative speakers, inferring what was really meant in an ungrammatical utterance and reformatting it. Most of these capabilities are still hard or impossible for computers. ”

As elaborated above, the meaning is the essential thing. Language is a data transmission channel to allow the creation of mental content in the mind of the listener as desired by the speaker. It is clear that the rules of grammar and language (word types, irregularities, categories, etc.) need to be learned (stored), however, language represents another level of quality for mental experiences in terms of perception, memory, emotions, and feelings. A lot of learning and personally evaluating of lower level mental content is involved before.

5.5 The Manual Dexterity of a 6–Year–Old Child

“At 6 years old, children can grasp objects they have not seen before; manipulate flexible objects in tasks like tying shoelaces; pick up flat, thin objects like playing cards or pieces of paper from a tabletop; and manipulate unknown objects in their pockets or in a bag into which they can’t see. Today’s robots can at most do any one of these things for some very particular object. ”

The addressed capabilities are closely related to imagination. Imagination is the reasoning about imaginary actions. It requires a large storage pool of already perceived actions and thoughts together with their evaluation in various respects. On the other hand, those capabilities require perfect motor control – which is something where machines can finally (in the context of this paper) compete with or even outperform humans.

5.6 The Social Understanding of an 8–Year–Old Child

“By the age of 8, a child can understand the difference between what he or she knows about a situation and what another person could have observed and

therefore could know. The child has what is called a “theory of the mind” of the other person. For example, suppose a child sees her mother placing a chocolate bar inside a drawer. The mother walks away, and the child’s brother comes and takes the chocolate. The child knows that in her mother’s mind the chocolate is still in the drawer. This ability requires a level of perception across many domains that no AI system has at the moment.”

Actually, this is the only point really in connection with psychoanalysis – as a theory about the adult mind. (see above), and being able to “mentalize” or having a “theory of mind” really is a crucially important capacity for any social interaction to succeed: we always consider the other’s point of view: what they might be thinking or feeling at the moment (always being aware of the fact that we will never be able to fully know what it is they are thinking or feeling), which mental states make them behave in a certain way, or considering that they do not know what we know. This capacity works in the background and is certainly not consciously intended, but still it is there and absolutely necessary for our understanding of ourselves, other persons and the interactions we come upon.

6 Necessary Basic Concepts

The above depicted road-map takes some concepts as granted. Among them are embodiment, a complex environment and a social system. This section gives a short overview on these topics.

Following the concept of “embodied cognitive science” as described in [32], intelligence must have a body that again shapes the way an agent is able to think. The body can be clearly distinguished between the surroundings, although the body has to contain an interface that grants bidirectional communication to the surroundings. Therefore, the body contains different sensors to sense the surroundings of the agent. Compared to the human’s body, the sensors take over the functionality of our five senses (taste, sound, tactile, sense of smell and vision) but have, due to the duty of the agent other functionality. To directly interact with the environment and fulfill the requirement of proactiveness of an autonomous agent, it also has to be equipped with actuators. The more complex the actuators are, the higher is the degree of freedom, the agent can interact and the more possibilities exist to reach desired goals. To lend importance to the body, it is equipped with internal sensors, that are monitoring internal values that are, together responsible for the homeostasis of the agent. This can be in terms of robotic agents e.g. energy level, processor load or internal network load, responsible for fast and slow internal message systems, comparable to the humans hormone system. Since, an agent shall be able to learn, also direct feedback through the internal sensors is eligible as a result of taken actions. Therefore, the environment has to contain dangerous components to gain deeper insight into its own body and allows approximating limits that are given in the environment. Each agent is placed into an ecological niche – defined by the environment and the possibilities of the agent to interact with it. According to Brooks [33], intelligence is within the environment and not within a problem solver. Thereafter, an agent is as intelligent as manifolded and/or complex its senses and actuators are.

An environment in which agents operate has to contain enough distinguishable objects and areas to enable the agent to navigate within it. Like watchdogs have good understanding where home and where outside is, software agents could then develop similar abilities. Also emotional cathexis of objects leads in a complex world to better results. If only few different obstacles are available, almost every place will have the same cathexis. The home will then be very positive, areas close to enemies are emotionally difficult.

The social understanding of an 8 year old as depicted above is only a small fraction of the complex social structures present in the human world. With added complexity and intelligence of (multi-) agent systems dynamic societies are needed. They are less complex as human ones, but nevertheless comparable to them. A society needs social rules (predefined or emergent). Social rules can be help others, don't go there, etc. they glue the society together. They are also some kind of reward system – he who owns great social reputation may receive more help and has more possibilities to influence others. Different societies have different characteristics: some value the experience of the elders, some the creative approaches, some strong leaders, etc. This is dependent on the needs of the society. A further advantage of societies and their social rules and social system is the possibility to specialize. One agent on its own has to do everything, in combination with others he can specialize into one task. This could also be done by design of a special task agent, but the specialization of an all purpose agent into a few tasks has the advantage that if other agents fail to deliver the required product/task (which has been agreed upon via the social system) the agent can do it by himself – albeit in lower quality or slower.

7 Implementation Issues

Implementing traditional AI demands for mathematical skills, using optimized algorithms and data structures. The above concepts, however, are very heterogeneous and have a rich structure. We have no reason to believe that an implicit architecture might be trained to show the desired behavior. Therefore the possibility of an explicit implementation shall be assumed here. Explicit in this context means that the various functional components and mechanisms, as described above, find their direct manifestation in a technical artifact – let's assume functional code.

Theoretically any technical function, described by a formal language, can be implemented solely in hardware – or entirely in software. This assumption holds, as long as we move withing the von-Neumann world of machines. Taken as a black box, no-one cares about the engineering principles inside, as long as it behaves on its interfaces as specified. Generating a specific mathematical function might be done via a lookup-table or via a formula, while both can be implemented and executed both in hardware or software (executed on an abstract, universal machine). The level of abstraction that the implementation shows towards the real meaning is sometimes called the “semantic gap”, discussed in [4]. Its consequence is decreased performance. The larger the gap, the worse the performance. This

derated performance can take extreme dimensions when the problem contains chaotic aspects. The human mind in general show extremely complex, nonlinear and chaotic behavior, stemming from incredibly large amounts of weighted memories that take influence on decisions. This exposed chaos gives an impression on the complexity of the underlying information processing. An implementation will face the same complexity and therefore strongly benefit from an as small as possible semantic gap. The following “software-near” hardware are examples for typical building blocks used in “small-gap machines”:

- Associative memory with database-like capabilities in hardware
- HW support for matrix operations
- Dynamically adaptive microcode/hardware
- Non-blocking multi-port memory
- Merging of memory and operations, hardware-in-memory
- Multidimensional, and not linear, connection of data and functions

All this is of course a violation of the pure von-Neumann idea of an abstract machine that outsources all complexity and functionality into time and software code. As current computers are (by implementing aspects of the above list) already far away from this pure idea, this philosophical aspect should not bother us.

In general, the above ideas of an intelligent machine demand a high level of parallelism in software and subsequently in hardware, if we want operations optimized. Parallelism, in turn, is a challenge to our technology, captivated in the boundaries of space and time. 3D-silicon might be one step more but the level of parallelism we need for instantaneously and successfully querying a tremendously large memory reaches beyond that.

Generally we can expect two sources of innovation that might bring us to the stage to implement our abstract concepts

- Technological innovation
- Structural innovation

Even nowadays and throughout history these two phenomena have – typically in alternating roles – brought technology forward. Just think of the development of computers where technological progress like relais, tubes, TTL, CMOS etc. were – usually when the current technology was on its limits – interlaced with structural innovations like parallelism, superscalar architectures or multi-cores. We can imagine massively parallel description languages that manifest in molecular nanotech and other currently unbelievable technologies to be the platform of tomorrow. Structuring the functional blocks of our models, however, will still be necessary. Taking X Billions of neurons does not create a human brain. It is the structural order which makes it special.

One open question is the scalability of this concept. Implicit AI typically has good scalability characteristics. A distributed genetic algorithm might for instance be scaled via its population and genome size. Similar things can be said of artificial neural networks, but as discussed in [3], neural networks do unfortunately not surprise us with unexpected smartness if we allow for growing.

The capabilities of ANNs do not really aggregate. If “more” is wanted, the ANN must show some macroscopical structure. The above concept’s structure is, however, macroscopically not scaling. It does not necessarily get twice as smart just if you introduce two semantic memories than just one. The size of the individual components, especially filter networks or memories, surely can grow if the given problem demands it.

Scalability plays a second role when thinking about really intelligent machines. “Singularity” evangelists (see the June 2008 issue of the IEEE Spectrum for an excellent pro-con discussion on this topic) base their positive expectations on Moore’s Law and the daily experienced ever accelerating progress of technology. Ironically, also critics use the economy of scale as argument. Nothing in nature (energy, matter, etc.) can grow forever, and exponential growth of progress is – in their view – a subjective and potentially false perception. Even if our “truly intelligent” machines might not be the paradise host for our consciousness, once it moves out of its current medium, they might be attributed conscious attributes.

8 Social Implications

The relation between truly intelligent machines and mankind is subject of numerous science fiction stories, the murderous and tragic ones dominating. Isaac Asimov created a series of such stories, funny enough it is always man who fails and causes trouble, and never machines who are described honest and innocent. We should not underestimate the mental abilities of man to cope with this challenge when complex machines are all around. We can expect a pretty sober reaction by people that grow up with mass media and ubiquitous communication. These things will have their place in our world. More interesting is the question what place we will have in their world?

Generally we should distinguish between three different encounters that our society might have in future:

- Artifacts that are equal to animals
- Artifacts that are equal to humans (or at least appear like that)
- Artifacts that are superior to humans (post-singularity beings)

The first one can already be observed. Ranging from “tamagotchis” to Sony’s Aibo robot-dog people bond to definitely inanimate tinboxes, more intelligent artificial animals will cause even more passionate behavior. Nothing wrong with that, if an artifact in deed has the intellect of a dog, why should it not receive the same respect as a real dog? Machines equal to us is more of a challenge. Phenomena like love, sympathy, racism or compassion might need a review if we live door-to-door with these creatures. And they might discuss the very same problems. The ultimate gain that we could get out of such a situation is to overcome all antique concepts of difference and realize that conscious beings are all equal. Anyway, we still did not encounter something entirely new. Throughout the history of our planet we met equal cultures and had to cope with it.

Meeting an outclassing entity, however, is something that no-body is prepared for. Antique mythology tells of gods which are equipped with extraordinary powers but are still helplessly extradited to very human emotions like love or anger. They can even be outsmarted sometimes, so they are by no means almighty. Modern religions paint a picture of a supreme creator, immune to our weaknesses, but captured in his own goodness and love for us. It is superior but not free. A mixture – an autonomous, superior being – threatens us. Would they exterminate us like we treat bugs and bacteria or would they explain us – finally – the true meaning of life? At least the question of social implications would not lay on our shoulders anymore, our superior creatures would ponder about that.

9 Conclusion and Outlook

Objects and devices get smarter every single day. It is a win-win situation for the users – who can get more services by machines – and manufacturers who always look for new products and markets. Their smartness today is shown in terms of function, usability, design, energy efficiency, sustainable product life cycle and the like. All of these are observable behavior, none of them can be directly assessed in terms of directly influences. With conscious machines on the other hand it would be possible to demand e.g. energy efficiency or security for children. Machine consciousness when developed following the template of human consciousness needs to show degrees of itself. It needs to be designed in the way it could potentially work in a real, living body in its mediator role between endogenous demands and the actual environment. These functions are elaborated by Freud and his successors who developed a functional model of the human mind. Another attribute of this model lies in the concept of the primary and secondary process and its implications for higher levels. Other humanities do not possess functional models or provide only behavioral models. However, psychoanalysis turned out to be applicable to surprisingly large extent into technical terms and in the future into systems.

The development of machine consciousness relies on many interdisciplinary findings presented above, whereby computer engineering and psychoanalysis will be the main contributors. Some of the requirements are already formulated in concepts, some still lack any idea for implementation. However, we tried to give the interested reader an impression of what has to be done to achieve machine consciousness in our view. We also have stressed the borders: that even very sophisticated solutions for human-like behaviour in terms of moving arms and feet, producing facial expressions, following speech dialogs, etc. do not contribute to making machines more conscious. The key for any stage of development of higher order mental content lies in subjective evaluation of lower level mental content. This is what children need to do from very soon after birth, and so conscious machines will have to. Some remarkable capabilities resulting from this developmental process in children have been presented.

There is already a community, founded with the origination of the ENF – the first international engineering and neuro-psychoanalysis forum [40]. Many international researchers took notice of that event and even more than 100 came to attend. Things start to come together. We see this as radically new approach in AI. Many other approaches to reach human-like intelligence failed because of – in retrospect – clearly visible lacks in theory or methodology. This one is unique in terms of understanding and applying proven theories of human consciousness into machines. You are welcome to join us and enhance our efforts with your valuable ideas.

References

1. Baars, B.J.: Some essential differences between consciousness and attention, perception, and working memory. *Consciousness and Cognition* 6, 363–371 (1997)
2. Damasio, A.R.: *The Feeling of What Happens: Body, Emotions and the Making of Consciousness*. Econ Ullstein List Verlag, GmbH (1999)
3. Barnard, E., Palensky, B., Palensky, P.: Towards Learning 2.0. In: *Proceedings of ICST IT-Revolutions 2008, Venice* (2008)
4. Palensky, P., Lorenz, B., Clarici, A.: Cognitive and Affective Automation: Machines Using the Psychoanalytic Model of the Human Mind. In: *Proceedings of First IEEE Engineering and Neuro-Psychoanalysis Forum, Vienna* (2007)
5. Arlow, J.A., Brenner, C.: *Psychoanalytic Concepts and the Structural Theory*. International Universities Press, New York (1964)
6. Bateman, A., Holmes, J.: *Introduction to Psychoanalysis – Contemporary Theory and Practice*. Routledge, London (1995)
7. Chalmers, D.: The Puzzle of Conscious Experience. *Scientific American*, 62–68 (December 1995)
8. Chalmers, D.: Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2(3), 200–219 (1995)
9. Dennett, D.: Who’s on First? Heterophenomenology Explained. *Journal of Consciousness Studies*, Special Issue: Trusting the Subject (Part1) 10(9-10), 19–30 (2003)
10. Sandler, J., Holder, A., Dare, C., Dreher, A.U.: *Freud’s Models of the Mind. An Introduction*. Karnac, London (1997)
11. Freud, S.: Formulations on the Two Principles of Mental Functioning. In: Strachey, J. (ed.&Trans.) *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 12, pp. 218–226. Hogarth Press, London (1911)
12. Freud, S.: The Ego and the Id. *Standard Edition XIX*, 109–121 (1923)
13. Freud, S.: The Unconscious. *Standard Edition XIV*, 166–204 (1915)
14. Freud, S.: *The Interpretation of Dreams*. *Standard Edition IV & V* (1900)
15. Freud, S.: Three essays on the theory of sexuality. *Standard Edition VII*, 135–243 (1905)
16. Freud, S.: Beyond the pleasure principle. *Standard Edition XVIII*, 7–64 (1920)
17. Solms, M., Turnbull, O.: *The Brain and the Inner World*. Karnac, London (2002)
18. Solms, M.: Freud returns. *Scientific American*, 56–62 (May 2004)

19. Solms, M.: What is the “mind”? A neuro-psychoanalytical approach. In: Dietrich, D., Zucker, G., Bruckner, D., Fodor, G. (eds.) *Simulating the mind, A technical, neuro-psychoanalytical approach*, Springer, NewYork (2008)
20. Viola, P., Jones, M.: Rapid Object Detection using a Boosted Cascade of Simple Features. In: *Conference on Computer Vision and Pattern Recognition* (2001)
21. Mattern, F.: Ubiquitous Computing: Schlaue Alltagsgegenstände - Die Vision von der Informatisierung des Alltags. In: *Bulletin EV/VSE*, vol. 19, pp. 9–13 (2004)
22. Hainich, R.R.: *The End of Hardware, A Novel approach to Augmented Reality*. BookSurge Publishing (2006)
23. Lindwer, M., Marculescu, D., Basten, T., Zimmermann, R., Marculescu, R., Jung, S., Cantatore, E.: Ambient Intelligence Visions and Achievements; Linking abstract ideas to real-world concepts. In: *Proceedings of the conference on Design, Automation and Test in Europe DATE 2003* (2003)
24. Endres, C., Butz, A., MacWilliams, A.: A Survey of Software Infrastructures and Frameworks for Ubiquitous Computing. *Mobile Information Systems Journal* 1(1) (2005)
25. Lagendijk, R.L.: The TU-Delft Research Program Ubiquitous Communications. In: *Proceedings of the Twenty-first Symposium on Information Theory in the Benelux*, pp. 33–44 (2000)
26. Roman, M., Hess, C.K., Cerqueira, R., Ranganathan, A., Campbell, R.H., Nahrstedt, K.: Gaia: A Middleware Infrastructure to Enable Active Spaces. *IEEE Pervasive Computing*, 74–83 (2002)
27. Mavrommati, I., Kameronas, A.: The evolution of objects into Hyper-objects. *Personal and Ubiquitous Computing* 7(1), 176–181 (2003)
28. Gellersen, H.-W., Schmidt, A., Beigl, M.: Multi-Sensor Context-Awareness in Mobile Devices and Smart Artefacts. *Mobile Networks and Applications* 7, 341–351 (2002)
29. Lohse, M., Slusallek, P.: Middleware Support for Seamless Multimedia Home Entertainment for Mobile Users and Heterogeneous Environments, 217–222 (2003)
30. Want, R., Schilit, B., Adams, N., Gold, R., Petersen, K., Ellis, J., Goldberg, D., Weiser, M.: The PARCTAB ubiquitous computing experiment. In: *Proceedings of the Fourth Workshop on Workstation Operating Systems* (1995)
31. Bion, W.: A theory of thinking. *International Journal of Psycho-Analysis* (43), 4–5 (1962)
32. Pfeifer, R., Scheier, C.: *Understanding Intelligence*. MIT Press, Cambridge (2001)
33. Brooks, R.: Intelligence without representation (47), 139–159 (1991)
34. Brooks, R.A.: I, Rodney Brooks, Am a Robot. *IEEE Spectrum* 06.08 (2008)
35. Alter, T.: Qualia. In: Nadel, L. (ed.) *Encyclopedia of Cognitive Science*, pp. 807–813. Macmillan Publishers Ltd., Basingstoke (2003)
36. Bringsjord, S.: The Zombie Attack on the Computational Conception of Mind. *Philosophy and Phenomenological Research* 59.1 (1997)
37. Holland, O (ed.): *Machine Consciousness*. Imprint Academic (2003)
38. Penrose, R.: *The emperor’s new mind*. Oxford University Press, Oxford (1989)
39. Fonagy, P., Target, M., Gergely, G., Jurist, E.L.: *Affect Regulation, Mentalization, and the Development of Self*, 1st edn. Other Press (2000)
40. Dietrich, D., Fodor, G., Zucker, G., Bruckner, D. (eds.): *Simulating the Mind*. Springer, Heidelberg (2008)
41. von Förster, H.: *Wissen und Gewissen: Versuch einer Brücke*, 7th edn. (2006)

42. Yowell, Y.: Return of the zombie - Neuropsychanalysis, consciousness, and the engineering of psychic functions. In: Dietrich, D., Zucker, G., Bruckner, D., Fodor, G. (eds.) *Simulating the mind, A technical, neuro-psychoanalytical approach*. Springer, NewYork (2008)
43. Beck, F., Eccles, J.C.: Quantum aspects of brain activity and the role of consciousness. In: *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, pp. 11357–11361 (1992)
44. Penrose, R., Hameroff, S.: Orchestrated objective reduction of quantum coherence in brain microtubules: The “orch OR” model for consciousness. *Mathematics and Computers in Simulation* 40, 453–480 (1996)