[16] O. H. Mowrer, *Learning Theory and Behavior*. New York: John Wiley, 1960.
[17] J. Vel´asquez, "Modeling emotion-based decision-making," in *Emotional and Intelligent: The Tangled Knot of Cognition*. D. Cañamero , 1998b, pp. 164–169.
[18] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge, MA: Cambridge University Press, 1988.
[19] A. Ortony, "On making believable emotional agents believable," in *Emotions in humans and artifacts*, R. Trappl and P. Petta, Eds. MIT Press, Cambridge, MA, 2003.
[20] W. S. Reilly, "Believable social and emotional agents," Ph.D. dissertation, School of Computer Science, Carnegie Mellon University, 1996.
[21] C. Elliott, "The affective Reasoner: A process model of emotions in a multi-agent system," Ph.D. dissertation, Northwestern University, Evanston, Illinois, 1992.
[22] C. Elliott, "Components of two-way emotion communication between humans and computers using a broad, rudimentary model of affect and personality," *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, vol. 1(2), pp. 16–30, 1994.
[23] D. Moffat and N. Frijda, "Where there's a will there's an agent," *Intelligent Agents: ECAI-94 Workshop on Agent Theories, Architectures, and Languages*, pp. 245–260, 1995.
[24] J. Gratch and S. Marsella, "A domain-independent framework for modeling emotion," *Journal of Cognitive Systems Research*, vol. 5(4), pp. 269–306, 2004.
[25] A. Sloman and M. Croucher, "Why robots will have emotions," in *Proceedings of the 7th Int. Joint Conference on Artificial Intelligence*, 1981, pp. 197–202.
[26] A. Sloman, R. Chrisley, and M. Scheutz, "The architectural basis of affective states and processes," in *Who Needs Emotions? – The Brain Meets the Machine*, M. Arbib and J.-M. Fellous, Eds. Oxford University Press, Oxford, New York, 2004.
[27] A. Sloman, "Beyond shallow models of emotion?" *International Quarterly of Cognitive Science*, vol. 2(1), pp. 177–198, 2001.
[28] A. Sloman, "What sort of control system is able to have a personality?" in *Creating Personalities for Synthetic Actors*, R. Trappl and P. Petta, Eds. Lecture notes in AI, Springer Berlin Heidelberg New York, 1997, pp. 166–208.
[29] G. M. Edelman, "Naturalizing consciousness: A theoretical framework," *PNAS*, vol. 100, pp. 5520–5524, 2003.
[30] K. Oatley and P. N. Johnson-Laird, "Towards a cognitive theory of emotions," *Cognition and Emotion*, vol. 1(1), pp. 29–50, 1987.
[31] S. E. Cross and E. Walker, "Dart: applying knowledge based planning and scheduling to crisis action planning," in *Intelligent scheduling*, M. Zweben and M. S. Fox, Eds. Morgan Kaufman, 1994, pp. 711–729.
[32] J. Searle, "Is the brain's mind a computer program?" *Scientific American*, vol. 262, pp. 26–31, 1990.

1.2 Considering a Technical Realization of a neuropsychoanalytical Model of the Mind - A Theoretical Framework

Dietmar Dietrich, Georg Fodor, Wolfgang Kastner and Mihaela Ulieru[16]

As foundation for a paradigm shift in artificial intelligence we propose a bionic model that encapsulates psychoanalytic principles of the human mind based on which we map Sigmund Freud's model of the "psychical apparatus" in combination with Luria's dynamic neuropsychology into a machine. Motivated by the first paper of this book which outlined the state-of-the-art in artificial intelligence we suggest future research directions and obstacles that need to be overcome when moving forward towards building conscious machines that will be even able to perceive and act on emotions and feelings. This paper outlines the motivation be-

[16] This work was supported by the HarrisonMcCain Foundation

hind our joint effort where scientists of the fields of psychology, pedagogy and psychoanalysis on the one hand, and engineers on the other hand are involved. As first outcome of this joint work, a model for a technical realization of a neuro-psychoanalytical model of the mind is presented. Ongoing activities and research results based on this model are shown in the following parts of this book.

1.2.1 Motivation

Today's automation systems demand for a high number of data points (sensors and actuators) and controllers (intelligent units[17]) to meet all the requirements of the underlying process [1], [2]. For this reason, they can no longer be based on standalone (central) systems, but have to be handled by a multitude of sub-systems leading to a distributed approach. Specific considerations for control units of (high dynamic) sub-processes (such as safety issues for drive control systems dedicated for airplanes or cars) are not within the focus of this paper. This work has its origin in the "Smart Kitchen Project" started in 1998. The initial idea was to perceive scenarios typically found in a kitchen and adequately react to dangerous situations. Special emphasis was put on the use of readily available technology (i.e., fieldbus systems for data collection, and databases for storing scenarios). Several people (e.g., [3]–[7]) contributed to answer the following two questions:

1) What technology can be used for a straightforward realization?
2) Where is basic research necessary?

The "Smart Kitchen Project" was followed by a European Union funded project called SENSE ("Smart Embedded Net- work of Sensing Entities"), which started 2006. Beyond this successor the team has started several European independent projects, presented in the following papers. From our point of view, it is mandatory to get neuro-psychoanalysts involved in our efforts at the cutting edge. Therefore the integration of such experts is a condition-sine-qua-non for all further work in this field [8].

In response to such needs we aim to develop a holistic model for automatic control of processes that tightly interact with human beings and their environment, such as robots that support, for example, persons suffering of dementia. Here, decisions concerning the overall context (e.g., "safe cooking") are necessary. Such decisions can, mainly for two reasons, not be obtained through the traditional way when defining control algorithms. Firstly, the involved parameters are numerous and can sometimes not even be described formally. The second reason is that the memory systems of traditional control systems are much too insufficient for the kind of control algorithms needed in this project.

Researchers from the communities of AI (Artificial Intelligence) and CI (Cognitive Intelligence) have taken a similar approach, as described in the previous pa-

---

[17] With the expression "intelligent" the authors mean in this case the technical definition and not the psychological meaning.

per [9] [18]. At the beginning they adapted the psychological principle of symbolization resulting in knowledge representation. Then, in a second phase, statistical thods and learning algorithms were applied [6]. During a third phase they concentrated on the term “embodying”, realizing that the human mind depends on an individual body (independent whether it is virtual or not). At that time, one goal was to implement a representation of the world into robots where data from the outside are captured by sensors and robots operate relying on their internal knowledge base. The current fourth phase can be characterized by the search for definitions of emotions and corresponding feelings, and a possible way to implement them.

A holistic model integrating emotions and feelings (terms as e.g. used by Damasio in [10]) into the technical realizations of intelligence is still not existent. We aim to include psycho- analytical models, by this starting a fifth phase. Our work is inspired by Sigmund Freud who was the first to develop a model of a “psychical apparatus” and its behavior. We aim to obtain a technical realization of the psychoanalytical model of the mind, thus following a bionic approach.

1.2.2 Premises

Within the last decade efforts have been intensified to correlate psychoanalytic models with modern neuroscientific concepts [11] (this intricate step has been appreciated and accepted.) The International Society of Neuro-Psychoanalysis was founded as a scientific society dedicated exclusively to this mission [12].

A further linkage between completely different scientific fields, namely neuropsychoanalysis and the engineering science, will possibly lead to “cultural” difficulties. Psychoanalysis is still facing strong (also politically formed) reservations in the scientific community. Moreover fundamental principles and approaches from AI and CI have to be re-considered for this project. Therefore, it seems necessary to define several premises for this endeavor.

- Premise 1: Eventually, all functions of the brain/psyche will be understood. In the long term a modeling of all functions will be possible [13].
- Premise 2: Science in general and specifically neuro- science have traditionally declared that subjectivity can- not be studied in a scientific way [14]. Indeed, attempting to study subjectivity one is confronted with the enormous challenge that subjective experience is directly accessible for the subject only and can never be directly observed and measured by an outside (objective) entity. Nevertheless we share the psychoanalysts’ opinion that those subjective processes, which were left out by the neurosciences, have an immense and crucial meaning for the understanding of mental life. We acknowledge the extraordinary difficulty of the scientific approach to subjectivity by the psychoanalytical method of indirect observation and interpretation. The resulting model of a psychical apparatus is in our opinion

[18] Following the demands of the Engineering and Neuropsychoanalysis Forum (ENF) three papers shall provide the basis for this endeavor.

still the best available model and shall serve as the base for further studies in this field. Thus, science has to take on the challenge.

- Premise 3: The final engineering model has to be cooperatively developed by engineers working together with experts in neurology, psychoanalysis, pedagogy and psychology. We have to consider that each community has its own culture, methods and ways of thinking, and make every effort to ease the inter-community communication difficulties. The scientific methods and concepts of the respective scientific worlds must be mutually acknowledged and respected. The task of the engineers has to be to study possibilities of simulating or even emulating the psychoanalytic models, and if successful, to find methods to put them into practice. The task of the experts in the neuro- psychoanalytic field must be to define, together with the engineers, a model, which satisfies the requirements of engineering (Fig. 1.2.1).

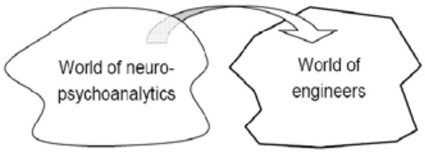

Fig. 1.2.1. Transfer of the models into the engineer's world

- Premise 4: If a complex organ, such as the psychical apparatus, is to be simulated or even emulated, a consistent model must be available. It is not acceptable to combine different descriptions of functions and types of behavior from various scientific approaches to the understanding of the mind (e.g., [15]) without evaluating the consistency of their combination.
- Premise 5: Sigmund Freud designed a functional model of the psychical apparatus with disregard to all anatomical and physiological correlates. He expected his contemporary colleagues to share this approach while working with or on the model. Similarly, the engineers working on this working with or on the model. Similarly, the engineers working on this project expect the psychoanalysts, in their work-groups, to form purely functional descriptions and to disregard the actual technological creation. It is quite evident that the engineers do not work with RNA, transmitters, and neurons but rather with silicon, transistors, controllers and computers. As much as the models rendered by the neurosciences are valuable inspirations for the actual technological realization, engineers would mostly refrain from any ambition to tackle the Sisyphus task of copying the biological systems.

1.2.3 Models

Algorithms used in AI and CI are based on symbolic logics and mathematical principles, respectively, which give them a key advantage: they are understandable and can – to some extent – be verified. Thus, contradictions can be excluded.

Experiments in psychoanalysis are not repeatable in a direct, trivial way. As already mentioned, subjectivity cannot be directly observed and reproduced in an objective way, although broad patterns can be extracted using statistics encapsulating the essence of several observations. Scientists have therefore decided to leave out subjectivity which they thought was not fit for scientific inquiry and opted for using statistical methods which allow distinct predictions. On this base specific behavior can be described in a clear way, but with the disadvantage of losing sight of the overall context. A unified and holistic model of the functionality of the mental processes, and how they work is still lacking. In the field of psychology and pedagogy many facets are explainable, but these facets cannot be put together to an overall model. For simulation or emulation an overall model is necessary.

The neurologist Sigmund Freud came to understand the high complexity and significance of the psychological processes that he found in his patients when focusing on their subjective experiences [16]. At the same time the neurological knowledge and the means of technical investigation of his time seemed to him to be much too insufficient to allow for a correlation of subjectively observed psychological processes with anatomical structures and physiological events in the brain. He therefore decided to observe the subjective psychological processes and interpret them in order to design a functional model of what he called the "psychical apparatus" completely disregarding the physical side of the equation. He named the so founded field of science "psychoanalysis". Departing from the existing schools of thought of his time, and specifically from the reductionist attempt of narrow localization, Freud defined instances of his apparatus like EGO and ID and their dynamic interplay. In his research work he analyzed the behavior of human beings and tried to explain their emotional and motivational aspects. Freud felt that the necessary correlation of his functional model of the psyche with neurological processes must be postponed until more knowledge and suitable technology was available.

Indeed, increasing efforts were launched to find such correlations. For example, A.R. Luria [17] developed his dynamic neuropsychology on the base of Freud's neuropsychological thoughts e.g., in his aphasiology. Luria's "objective" research is accepted and held in high regard in the neuroscience community.

Luria's work is of high importance for the integration of psychoanalytic and modern neuroscientific concepts specifically focused on by M. Solms [18], [19]. The results of these activities are of crucial importance for the authors when trying to bridge the gap between the research fields of the sensor and actuator areas and the "higher functions". If the idea is to develop a technical concept, this part is an important component for the whole chain of units.

If we are ready to accept these ideas, we have precise constraints which allow tackling the realization of the assumed model[19]. We aim to describe the brain with only one consistent model integrating three models.

---

[19] It is necessary to point out again that this paper of the ENF should not explain the whole context

Model 1: Neurological model (as a base for a model of communication, information flow and simple control functions of the human body)

The central nervous system can be roughly differentiated into two units, the brain and the peripheral nervous system. The brain, i.e., the "master station", is, from the point of view of computer technology, totally decentralized. Each neuron can be seen as an autonomous controller. The peripheral nervous system, connecting sensors, actuators and the brain, has communicating, but to a limited extent also computational functions.

To get an understanding of the operating principles of the brain, neurologists and biologists analyzed the topological structures and networking. Their results were previously the base for AI to design neural networks. However, this method offers only a limited chance for huge complex systems, a serious drawback if we realize that a human brain has billions of neurons. In the area of biology it is an efficient way to understand the information system of animals like bugs, flies or worms, which have only a small mass of neurons. Nevertheless, it is illusionary for the not so distant future to hope that scientists will be able to understand the connection between hardware and the higher functions of a brain like consciousness, if they follow this way of thinking – regardless of wild speculations.

The neurological knowledge can serve as model, to understand the lower level of the brain – if we assume a hierarchical brain system for the lower levels [20].

Model 2: Psychoanalytical model (as a base for a functional model of the psychical apparatus)

As mentioned above neurologists such as Freud and Luria recognized very early [17], [21] that the higher brain functions must be modeled as functional, dynamic systems. The psychoanalytic theory is based on the idea of a psychical apparatus being a functional system but as mentioned before contains no models of anatomical systems or physiological processes correlating with the mental processes.

A considerable gap between both models opens up which makes a further model necessary to bind model 1 and 2.

Model 3: Link between the neuron system and the psychoanalytical model (between model 1 and 2)

A decisive point in modeling the mental system is the correlation between the psychological processes and the physiological processes of the nervous system. Knowledge about this will considerably increase our understanding of the mind.

We will adhere to Freud's idea of this relationship which is fundamental to his entire psychoanalytical concept. The fundamental proposition, so elaborately discussed by Solms [22] is that mental processes are in themselves unconscious. Consciousness is a mere reflection or perception of mental activity. The psychical apparatus has two perceptual surfaces generating the totality of conscious experience: One surface is directed towards external objects and the processes they are

---

and all the details, which were worked out up to day. The goal is only to present the idea of the new research step and the vision of it.

involved in representing the existence of things including our physical body and proprioception. The other surface is directed towards the inside perceiving psychical states that represent processes occurring inside ourselves.

Consciousness serves the perception of both classes of sensory input. These two classes are registered on two different perceptual surfaces, one facing outwards, the other one inward. These surfaces are hierarchically equivalent. One does not produce the other. They have rather qualitatively different ways of registering reality, which, as I. Kant reminds us, is and will always remain unknowable in itself. Therefore what we perceive as a physical object on the outside – the brain – appears as a subjective psychical world viewed from the inside; one and the same thing – the psychical apparatus – perceived from two different perceptual surfaces (cf. [23], p. 141).

Based on this proposition the authors of [24] have tried to find a first bridging theory. Today this approach must be tackled in a more differentiated and modified manner.

In the field of computer technology a connection between hardware and software is relatively easy to define unambiguously, if one only considers the aspect of the system description. It means that this part of the system can be defined as a link (or interface) if it is possible to describe this part as hardware as well as software (both are abstract formalisms). The micro program control unit of the microprocessor can be regarded as such a part. This unit – because it is pure hardware – can be described by a hardware description language. On the other hand – because of efficiency reasons – the program of this unit is nowadays usually described in a micro program language, which is specific for the respective microprocessor. So, in this way the micro program unit represents the link between hardware and software.

The micro program unit is the complex control unit for the microprocessor, and represents the base on which the higher level software is mounted. These function levels (above the micro program) are the drivers, which are part of the operating system, the operating system itself and the application units[20] of the computer.

Beside the hierarchical configuration of the functional units in a computer system, the different languages can also be ordered in a hierarchical way. The machine respectively the assembler language is situated above the micro program. All of them are hardware specific. Above them the high languages are defined, and in the next level the functional languages [25], [26].

According to [11], [17], [27], one can regard the lower functions of the brain as a system of (abstract) hierarchical levels. The cortical regions can be differentiated into three areas (cf. Fig. 1.2.2): Luria defined the primary region as the *projection field*, which receives data from sensors, and sends commands to actuators; for higher order processing, the association *field* and as a next higher level the *comprehensive regions*. This classification may help to describe different levels of ab-

---

[20] Application units are systems like word processing applications or programs which control machines.

straction and integration from neurological units to higher functions in a way analogous to how computer systems are designed[21].

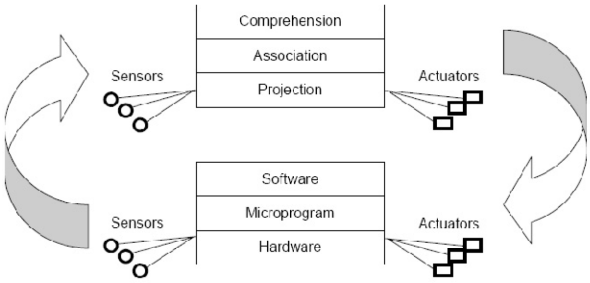

Fig. 1.2.2 Functional models for transfer and computation of data in both worlds with different levels of abstraction

Hardware and software are functional units combined by a hardware/software interface (micro program). In order to allow matching the computer model and the model of a human mind, both models have to be defined in a transparent and strongly modularized way. However, we have to keep in mind, that we will never be able to compare each level 1:1. The unified view on both models must be comprehensible for all parties involved, the engineers and the neuro-psychoanalysts.

Our proposal is to use symbolization as the interface between model 1 and model 2 which will be explained in detail in the next section.

1.2.4 Symbolization of the outer world

Invoking ideas of [27]–[29] a model which was already partly presented in [1], [24] is useful. Human beings' actions are based on experiences and their own behavior learned in past scenarios. The infant acquires knowledge of the outer world by learning processes. This means that the embryo and infant cannot perceive raw data from the outside world. The flood of data coming from all its sensors is initially extensively diffused. The infant has to learn to transform perceptions of outside objects and the processes they are involved in into symbols of rising levels of abstraction (Fig. 1.2.3). The representation of the Self and the outside world is increasingly composed by the process of symbolization (in the projection field) producing integrated images (in the association field). Thus, two function units can be differentiated: one unit where all symbolized objects are memorized and the representation unit which will be described in more detail in the next chapter.

To understand Fig. 1.2.3 one has to consider that the output of the eye[22] (after the neuron layers in the eye) are not pixels like a camera, but only characteristic

---

[21] We have to consider that computer systems are described in such an accurate way because the designer possesses tools for all different abstraction and language levels. The computer expert usually synthesizes systems and does not analyze them. The neurologist and psychoanalyst try to understand nature which means they have to go the opposite way which is incomparably more difficult.

[22] The following statements are valid for all senses.

values like areas, edges or arcs, which the brain combines to form images. Images of the inner world may thus be regarded as a matrix (collection) of symbols [29], which are again assembled in the representation unit.

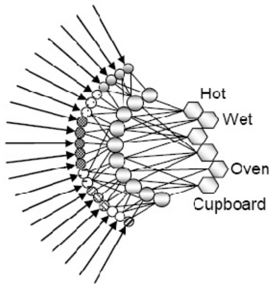

Fig. 1.2.3 The process of symbolization: condensation of sensory input to higher order symbols

According to [28] it was an immense achievement to find out that the embryo and the infant first have to learn, from the diverse, incoming dataflow, to compose all the images. In this case we do not only speak about optical and acoustic stimuli, but also about images of the olfactory organ, sense of touch and images of the motion of the own muscles. The embryo and infant are not able to differentiate between the physical inner and outer world. In the beginning they only understand one holistic image of the inner and outer world as one whole object. The perception of objects and their dynamic behavior are a computational "work" of associated images by an incoming data flow and the data from memory (Fig. 1.2.4). The human being "sees" a virtual image which is the result of a complex computational neural process of matching incoming data against stored information/knowledge.

We refer to the representation process, including the association of images as "the image handling" (in the association field). It can be regarded as the base for all higher functions of the mental process which are described by the psychoanalytical model. How this image handling could work on the base of neural networks has to be investigated. However, this is not within the scope of this article.

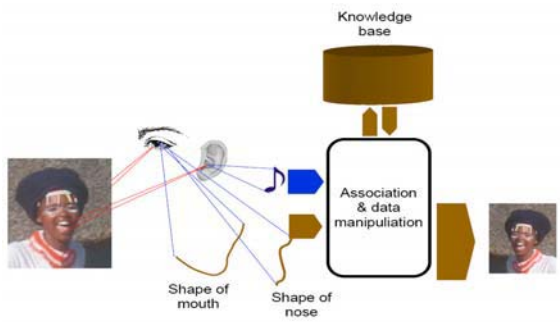

Fig. 1.2.4 Assembling of images in the inner world [30]

It is important to differentiate between "perception" and "recognition". Perception uses symbolization only. In contrast to human beings, bugs and worms only have few neurons. Therefore, it seems to be reasonable to assume that these creatures rely on (simple) symbolization, which is also a principle of the lower level functions of human beings. These lower level functions can be described by means of mathematical algorithms such as fuzzy logic algorithms or statistical methods at the sensors level, and with symbolization above it [6], [7]. For the association of the pictures AI offers different procedures, which can be partly solved using hardware and/or software. To find optimal solutions for this will require much effort and will still remain a difficult task. In our model "recognition", in contrast to perception, involves feelings.

1.2.5 Extended model: association and projection fields – emotions and feelings

The models of the psychoanalysts and their modularization of the psychical apparatus may be sufficient for their work, and it was also the base for our first research steps. However, we had to realize soon that the functional units, as they will be described in the following papers of this forum, were not differentiated and distinct enough for a technical mapping because the single units were too complex for a clear technical definition. Psychoanalysts have solved the problem for themselves by using different models, which they defined from different point of views. As the engineers are only able to work with one *unified model*, their task is now to look for further concepts. These concepts should further refine modularization of the psychoanalytic entities without contradicting them [31] [23].

Solms differentiates between simple or primary consciousness (PC) and reflexive or extended consciousness (EC) (cf. [11], p. 95). This hierarchical concept is brilliant and extends Luria's structure. The primary consciousness fulfils the requirements of Luria's projection and partly association field. The extended consciousness corresponds to Luria's comprehensive field. All information, which is

[23] We believe that mixing several models will not lead to a feasible solution. Although interesting. for further technical realizations, such "mixed" modeling would bring to great confusions from the psychoanalytical perspective.

captured by sensors, will be supplied to the primary consciousness. Perception is in this meaning the presentation of sensory data condensed by a symbolization process like the symbolic information wet or *hot* (cf. Fig. 1.2.3) or like a data field of characteristic values of an optical or an acoustical image of the outer world like the object *wardrobe*. Images of the inner and outer world are composed by symbols. The different images of the optical, acoustical or olfactory channels are in such a way pure logically and mathematically computed data, composed by symbols of higher abstraction levels (Fig. 1.2.5). Scenarios are short sequences of images and will also be memorized in the brain like the images. A much higher level of intelligence, the comprehensive field [5], is necessary for acts which are represented in sequences of scenarios, which composing must be a much more complex function.

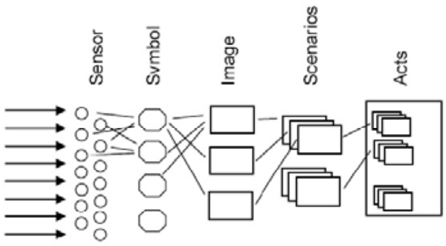

Fig. 1.2.5 Abstraction levels for computed data

Furthermore, a representation field (representation layer), which is the perception unit of intelligent creatures (as explained in the previous section), is situated in the association field. This unit gets data from the outer world and elicit the association of images and scenarios at the same time which were developed in the past. Both these input resources, the channel with data coming from outside and the data associated, are combined and develop these images and scenarios which we believe to perceive. The process can be seen as a pure mathematical procedure. The result is assembled unconsciously. What we perceive is therefore not the optical picture, passing the lens of the eye, or the sounds, passing the eardrum, but an image of the world which the intelligent subject has developed by means of a complex procedure with different kinds of data (Fig. 1.2.4). The science of today is not able to say how much percent of the input is coming from outside and how much from the database. We assume that the biggest part of these images comes from the "knowledge base" of our brain, because the throughput of sensor data is poor in opposite to all the images and scenarios which we are able to see in a fast sequence.

As mentioned before, insects only have a few neurons in contrast to human beings, and similarly perhaps a projection or association field but never a comprehensive field. They react in a purely "mechanical" way, similar to our concept of a robot. This means in the language of electrical engineering that only control loops (mathematical models) and/or if-then-rules (logical reasoning) are the base of their intelligence. Damasio wrote in [32], that a typical representative of such an operating mode was Mr. Spock a well-known science fiction character from the Star Treck series. It is interesting that he was presented to be superior to human beings. “He was (nearly) not influenced by feelings.” In contradiction to this Damasio wrote, that nature – in a Darwinistic view – tells us exactly the opposite. Feelings are the base for the comprehensive field and therefore the base for a high intelligence. The behaviour of Mr. Spock is only understandable in a mathematic/logical reasoning. In comparison to the human beings’ Spock’s performance is much poorer. A human being is capable of much more which turns out to be a selection advantage.

Nature developed the comprehensive field. For the evolution of human beings it was a very important step to set the comprehensive field above the projection and association field which allows the formation of the Self [11].

When constructing a hierarchical resolution, the reflex arc could be placed as a simple control loop into the lowest “intelligence level”, as Norbert Wiener has already described it [33]. The projection and association field could be placed into the next higher level, the comprehensive field into the highest intelligence level. It is obvious that especially the upper two levels have to be differentiated and further modularized into sub-levels and sub modules. The comprehensive field is even more complex than the projection field [34]. The core functions of the projection and association fields, namely representation, memorizing of images and scenarios, and symbolization occur in the comprehensive field too. However, the decisive aspect is that the comprehensive field needs to be constructed with at least two representative layers, which means that in our approach the mental system will include three representation layers at least: the first being the emotional representative layer in the association field, next the representation of the outer world, and finally a representation of the Self. Thus, consciousness also means that the subject sets itself in relationship with the outer world. The human being is seeing itself as a person vis-a-vis of the outer world. He or she is something distinct from it. Because the representation of the Self can only be developed by images, it is easy to assume that the Self is again nothing else but a vast collection of images and scenarios [35], and the representation of the Self at any given moment is only a short snapshot of a huge number of various images and scenarios permanently changing. [24] These images and scenarios can be associated “all the time” from our knowledge base, triggered by symbols, coming from the inside world but also formed by data coming from outside. This is what makes the human being so difficult to describe. He cannot be represented by only one image or a single algorithm. A lot of contradictory images can be memorized. The image of the Self is specifically formed by images and scenarios, which were laid down as memories in the beginning of a human being’s life and are never forgotten [11].

[24] Here it becomes understandable that because of the huge number of memorized various outer worlds and the Self images and scenarios, experiments with human beings are not repeatable like physical experiments.

The representation of the Self has two disadvantages, one being that the images are from the past all the time (in contrast to the outer world) and the other, that it is very difficult to superimpose them with newer ones. There is always a difference in how one sees oneself compared to reality. This poses a special problem for engineers during the phase of implementation.

To be able to take the first step into the direction of a psychoanalytic inspired bionic system we will start from the question: How can feelings be defined in contrast to emotions and how do emotions work in the comprehensive field [36]? As explained above, the research team in Vienna [5] defined emotions as symbolized data (in the projection and partly in the association field). They are value-free and inform the "process", representing the creature, about physical states and behaviour. The snail senses whether it is wet outside, the fly senses whether the air flow is rough.

Feelings, as opposed to emotions, are valuations. In the comprehensive field, symbols connect feelings with valuated images and scenarios. Symbols, images and scenarios are memorized in a "weighted" way. If a symbol is formed, it associates not only a reaction image in the projection and association field but also images from the inner world, which are images from the past. They are evaluated and then linked with the Self (a term as used in psychoanalytic theory) creating a new feeling, which is a complex cumulative value. Depending on the current inner state an input from outside initiates a particular feeling which depends on the knowledge of the past. This is why repeatability in experience with the human psyche is so hard to obtain.

With these notions it is also possible to differentiate between perception and recognition. The Vienna team applies the definition that perception (achieved by sensing) is situated on the level of projection and association field and recognition on the level of comprehensive field and this has something to do with feelings.

1.2.6 Open questions and proposal for technical realization

Three representation layers have been identified by the Vienna team (see 1.2.5): one for the projection and association field and two for the comprehensive field. It is a hypothetical model and needs to be proven. The idea was: Following our definition that the projection and association fields do not include feelings, the representation layer for both fields can be seen as a simple architecture [3], [5]. Incoming symbols are classified according to elementary sensations, e.g., colour, brightness, loudness level or heat [37]. After a first "computation" – for fast reactions – they can trigger an action depending on the scenario which is recognized (which also means that the outside scenario must be similar to the memorized scenario). For these steps, not so complex algorithms can be applied [3].

The upper representation layers are much more complicated. Feelings are involved. The scientific literature of neuro- psychoanalysts does not provide an answer clear enough for the kind of model engineers need for technical realization. We know that the tasks of these representation fields are to take care for what we can "see" or "smell". We engineers understand that these representation fields play a decisive role for the human being's consciousness. One representation layer is responsible to "see" the Self, the other for the physical body in the outside world (Fig. 1.2.6). In this sense the Self is a virtual person, an understanding of oneself, strongly influenced by the own homoeostasis. However, there are a lot of upcoming questions, which cannot be answered at present. How are the images and scenarios for the representation layer of the Self composed? What stands behind the symbol "Σ" in Fig. 1.2.6? Must be differentiated between emotions and feelings in the model? What is a representation field? Solms writes in [11]: "If the brain is dreaming, the data channel coming from the sensors are turned back to the knowledge base, and the brain initiate itself to deliver images and scenarios to the representation field". Who determines the first images and scenarios? What affects the course of the dream? The answers to these questions psychoanalysts can provide are not sufficient for technical realization.

We have to find solutions for all these questions to be able to synthesize the model with the different facets.

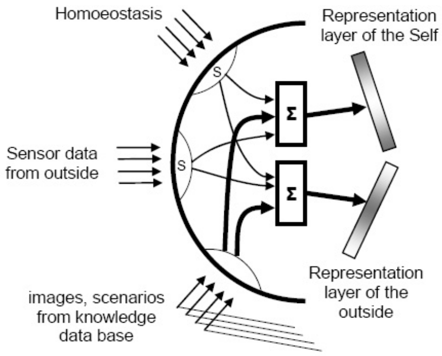

Fig. 1.2.6 The two different representation layers of the comprehensive field (S: Symbolization)

Excluded issues

More difficult topics like learning, forgetting, sex-specific differences and sexuality are not addressed for the moment. Today, the team concentrates on the simpler aspects of the psychoanalytic models, which seem to be technically solvable.

Constraints of the model

Following a bionic way of thinking, we as engineers try to emulate the architectures found in nature. However, it is not within our goal to copy them. As such it is definitely not the goal of our research work to copy the human being. We consider this issue to be a matter of philosophical and ethical concern.

1.2.7 Conclusion

The present project has the goal to adapt, simulate and to emulate parts of the psychoanalytic model for bionics applications. If we try to formulate and prove all upcoming questions carefully, then a possibility to map the principles into the world of engineers is realistic.

Currently, control systems lack the possibility to perceive complex scenarios. Damasio's idea that the development of a comprehensive field was a necessary step is consistent with Darwin's principles [10]. So far, engineers have used bionic approaches. Also, AI was successful in taking over neurological principles. Starting from the current state of the art we aim to go one step further and analyze higher functions of the brain together with neuro-psychoanalysts. In anticipate that this will help to describe complex scenarios for all kinds of automation systems in a better way.

Besides automation systems, we also expect the applicability of our model for various other application fields. Take the long standing research field of speech recognition as an example. Often, research institutes have announced a breakthrough, but each time the real success was more than modest. Elaborate algorithms based on semantic rules proved insufficient to tackle this complex problem. Thus, the success story of speech recognition systems is still limited, although they are applied in restricted domains (e.g., medicine). A universally applicable machine which understands sentences from independent speakers does not exist to date. We would like to support a radical shift in approaching complex problems such as speech recognition by bringing to the attention of the engineering community works overlooked from the field of psycho-analysis, such as Sigmund Freud's "Zur Auffassung der Aphasien – eine kritische Studie" [21] in which he criticised 100 years ago the neurological models of that time. Without question, neurology has made big progress up to date. However, we believe that if engineers read this fundamental paper of Sigmund Freud they would have considered the speech recognition problem in a different light which involves feelings and perception as delineated in this paper.

We claim this would have been a better starting point not only in the development of speech recognition systems.

References

[1] D. Dietrich and T. Sauter, "Evolution potentials for fieldbus systems," in *Proc. IEEE International Workshop on Factory Communication Systems (WFCS'00)*, Porto, Portugal, Sept. 2000, pp. 343–350.

[2] G. Pratl, W. Penzhorn, D. Dietrich, and W. Burstaller, "Perceptive aware- ness in building automation," in *Proc. IEEE International Conference on Computational Cybernetics (ICCC)'05*, Port Luis, Mauritius, Apr. 2005, pp. 259–264.

[3] G. Russ, "Situation-dependent behavior in building automation," Ph.D. dissertation, Univ. of Technology Vienna, Vienna, 2003.

[4] C. Tamarit-Fuertes, "Automation system perception," Ph.D. dissertation, Univ. of Technology Vienna, Vienna, 2003.

[5] G. Pratl, "Processing and symbolization of ambient sensor data," Ph.D. dissertation, Univ. of Technology Vienna, Vienna, 2006.

[6] D. Bruckner, "Probabilistic models in building automation – recognizing scnearios with statistical methods," Ph.D. dissertation, Univ. of Technology Vienna, Vienna, 2007.

[7] W. Burgstaller, "Interpretation of situations in buildings," Ph.D. dissertation, Univ. of Technology Vienna, Vienna, expected May, 2007.

[8] The Artificial Intelligence Recognition System (ARS) Project. [Online]. (2007) Available: http://ars.ict.tuwien.ac.at/
[9] B. Lorenz and E. Barnard, "Artificial intelligence – paradigms and applications," in *Proc. 1*st *International Engineering & Neuro- Psychoanalysis Forum (ENF'07)*, Vienna, Austria, July 2007, p. in print.
[10] Damasio, *The Feeling of What Happens, Body and Emotion in the Making of Consciousness.* New York, NY: Harcourt Trade Publishers, 1999.
[11] M. Solms and O. Turnbull, *The Brain and the Inner World: An Introduction to the Neuroscience of Subjective Experience*. New York, NY: Karnac Books Ltd., 2003.
[12] The International Neuro-Psychoanalysis Centre and Society. [Online]. (2007) Available: http://www.neuro-psa.org.uk/
[13] W. Freeman, *Society of Brains – A Study in the Neuroscience of Love and Hate*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1995.
[14] W. Balzer, *Die Wissenschaft und ihre Methoden. Grundbegriffe der Wissenschaftstheorie*. Munich, Germany: Alber Verlag, 1997.
[15] C. L. Breazeal, *Designing Sociable Robots*. Cambridge, MA: The MIT Press, 2002.
[16] M. Solms, "Freud, luria, and the clinical method," in *Psychoanal. and History*, 2000.
[17] R. Luria, *Working Brain: An Introduction to Neuropsychology*. London, UK: Penguin Books Ltd., 1973.
[18] M. Solms and K. Kaplan-Solms, *Clinical Studies in Neuropsychoanalysis*. London, UK: Karnac Books Ltd., 2000.
[19] M. Solms, "*The Neuropsychology of Dreams – A Clinico-Anatomical Study (Institute for Research in Behavioral Neuroscience)"*, Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1998.
[20] G. Pratl and P. Palensky, "The project ars – the next step towards an intelligent environment," in *Proc. IEE International Conference on Intelliegent Environments*, Essex, UK, June 2005, pp. 55–62.
[21] P. Vogel, "*Sigmund Freud zur Auffassung der Aphasien – eine kritische Studie (2. Auflage)"*, Frankfurt am Main, Germany: Fischer Taschenbuchverlag, 2001.
[22] M. Solms, "What is consciousness," *Journal of the American Psychoanalytic Association*, vol. 45/3, pp. 681–703, 1997.
[23] S. Freud, "*An Outline of Psychoanalysis*", London, UK: Hogarth Press, 1940.
[24] D. Dietrich, W. Kastner, and H. Schweinzer, "Wahrnehmungsbewusstsein in der Automation - ein neuer bionischer Denkansatz," *at*, vol. 52, pp. 107–116, Mar. 2004.
[25] J. P. Hayes, "*Computer Architecture and Organization (2*nd *ed.)*", New York, NY: McGraw-Hill Publishing Company, 1988.
[26] Olsen, O. Faergemand, B. Moller-Pedersen, R. Reed, and J. Smith, "*Systems Engineering Using SDL92"*, Amsterdam, Netherlands: Elsevier, 1994.
[27] O. Sacks, "*The Man Who Mistook His Wife for a Hat"*, New York, NY: Simon & Schuster, 1987.
[28] M. Dornes, „*Der kompetente Säugling - Die präverbale Entwicklung des Menschen"*. Frankfurt am Main, Germany: Fischer Taschenbuch Verlag, 2001.
[29] H. Förster, „*Wissen und Gewissen*" Frankfurt am Main, Germany: Suhrkamp Taschenbuch Wissenschaft, 1993.
[30] E. Brainin, D. Dietrich, W. Kastner, P. Palensky, and C. Rösener,¨ "Neuro- bionic architecture of automation – obstacles and challenges," in *Proc. IEEE Africon*, Gaborone, Botswana, Sept. 2004, pp. 1219–1222.
[31] C. Rösener,¨ B. Lorenz, K. Vock, and G. Fodor, "Emotional behavior arbitration for automation and robotic systems," *in Proc. IEEE Inter- national Conference on Industrial Informatics (INDIN'06), Singapore*, Singapore, Aug. 2006, pp. 423–428.
[32] Damasio, "*Looking for Spinoza; Joy, Sorrow and the Feeling Brain"*. New York, NY: Harcourt Trade Publishers, 2003.
[33] N. Wiener, "*Cybernetics, or Control and Communication in the Animal and Machine"*. Cambridge, MA: The Massachusetts Institute of Technology, 1948.
[34] B. Palensky-Lorenz, "A neuro-psychoanalytically inspired cognitive architecture for autonomous systems," Ph.D. dissertation, Univ. of Technology Vienna, Vienna, expected April, 2007.

[35] . T. Deutsch, R. Lang, G. Pratl, E. Brainin, and S. Teicher, "Applying psychoanalytical and neuro-scientific models to automation," in *Proc. International Conference on Intelliegent Environments*, Athens, Greece, July 2006, pp. 111–118
[36] W. Burgstaller, R. Lang, and P. Pörscht, "Technical model for basic and complex emotions," in *Proc. IEEE International Conference on Industrial Informatics (INDIN'07)*, Vienna, Austria, July 2007, p. submitted.
[37] G. Pratl, "A *bionic approach to artificial perception and representation in factory automation," in Proc. IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'05)*, Catania, Italy, Sept. 2005, pp. 251–254.

1.3 What is the "mind"? A neuro-psychoanalytical approach

Mark Solms

The brain is the object of the neurological sciences. The object of psychological science is the mind. Few people would disagree that the mind and the brain are ontologically indistinguishable. This begs the question: what is the 'mind' and how does it differ from the brain? In my view, the mind is distinguishable from the brain only in terms of observational perspective: the mind is the brain perceived subjectively. *Psychoanalysis is a branch of psychology that has taken this perspective seriously.*

Psychoanalytic study of subjective experience has resulted in a model of the mind which can be reduced to five components. (1) The driving principle of life is survival in the service of reproduction. (2) The function of the mind is to register survival/reproductive needs and satisfy them in the world. (3) Since the same could be said of the brain, the mind comes into its own by registering such satisfactions through feelings*. Feelings – pleasures and unpleasures – register the brain's biological successes and failures. This is the basis of consciousness. (4) Feelings generate the values from which* intentionality *is derived. Intentions boil down to wishes to repeat previous pleasurable experiences. This requires memory. (5) Experience, registered in memory, demands increasingly complex decisions about how pleasures can be obtained in reality. This in turn demands response flexibility, which is achieved through thinking. Thinking is experimental action. It depends fundamentally upon response inhibition. This is the basis of 'agency'. Agency is the freedom* not *to act.*

Attempts to manufacture artificial minds must replicate these functional principles.

1.3.1 Introduction

Since our engineering 'colleagues' ultimate aim seems to be the construction of an artificial mind – and since they wish to use our (neuro-psychoanalytic) knowledge in this regard – it is an ideal opportunity to address the question I have framed in my title: *what* is a 'mind'? In the process of addressing this question I will of necessity also consider two related questions: *where* do minds occur in nature? (Localization); and *why* do they exist? (Function)

2) It is one thing to address such questions, and another to do so neuro-*psychoanalytically*. What is special about the neuro-psychoanalytical ap-