One of the most dramatic milestones in the evolution of animal life is the emergence of an outer skeleton, operated by 'voluntary' striated muscles, in the lancet fish, the first vertebrate animal. This innovation resets the stage for the interior-exterior problem altogether. Indeed, as Mark Solms indicated in his talk at the ENF, it might be argued that in vertebrates it is as if two bodies, each with their proper cohesion, are continuously operating in parallel: an 'internal' body, inherited from the invertebrates and where smooth muscles and glands are the predominant effectors, and an 'external' body, emerging in the vertebrates and where the 'new' voluntary striated muscles, operating the outer skeleton, are the predominant effectors. The internal body is the source of major stimulus production – such as respiratory needs, hunger, thirst, sexual tensions etc. – which threaten the stability of the organism. This internal body, thereby, produces demands for work upon the external body: the external body is *driven* to interact with the external world so as to ultimately make the internal stimulation stop. At the same time, the external body is also a source of major stimulus production – namely those arising from the perception of the outside world. It thus probably has to manage its own stability in terms of keeping the potential energy at a non-threatening low level. It is as if a vertebrate's mind, then, faces the problem of having to manage two stimuli-metabolisms, each with their proper inner-outer complexity, as well as to attune both metabolisms reciprocally (i.e. the inner with the outer body and vice versa). I agree with Mark Solms that consciousness might critically serve for these multiple adjustments of the interior-exterior problem.

References

Bazan A (2007). An attempt towards an integrative comparison of psychoanalytical and sensorimotor control theories of action. *Attention and Performance XXII*. Oxford University Press, New York, pp. 319-338.

Bazan A, Van de Vijver G. (*in press*). L'objet d'une science neuro-psychanalytique. Questions épistémologiques et mise à l'épreuve. In: L. Ouss, B. Golse (éds.), *La Neuro-Psychanalyse*. Odile Jacob.

Blakemore,S.- J., Frith, C. D. and Wolpert, D. M. (1999). Spatio-Temporal Prediction Modulates the Perception of Self-ProducedStimuli *Journal of Cognitive Neuroscience* 11: 551-559.

Freud, S. (1895/1950). *Entwurf einer Psychologie*. G.W., Nachtragsband, pp. 375-477.

Freud, S. (1895/1966) Project for a scientific psychology (Stratchey, J., trans.). In Standard Edition I, pp. 281–397/410. Hogarth Press, London. (Original publication in 1950.)

Freud, S. (1915). Instincts and their vicissitudes. *Standard Edition* XIV: 117-140.

Helmholtz, H. (1878/1971) The facts of perception. In Kahl, R. (ed.), *Selected Writings of Hermann von Helmholtz*.Wesleyan University Press,Middletown, CT.

Lenay, C. (2006) Enaction, externalisme et suppléance perceptive. *Intellectica*, 43, 27–52.

Shevrin, H. (1998) Why do we need to be conscious? A psychoanalytic answer. In Barone, D. F. Hersen, M. and Van Hasselt,V. B. (eds), Advanced Personality, Chapter 10. Plenum Press, New York.

Stainier, D.Y.R. (2005). No organ left behind: tales of gut development and evolution. *Science* 307: 1902-1904.

3 The Mental Apparatus for Complex Automation Systems

A Combined Computer Scientific and Neuropsychoanalytical Approach

Dietmar Dietrich, Mihaela Ulieru[60], Dietmar Bruckner, Georg Fodor

The INDIN/ENF [ENF 07] heralded the start of a paradigm shift in the design of intelligent systems. As the emerging community of interest further investigates ways to move forward on this path, we herewith underline how, if several boundary conditions are kept, the hypotheses and suggestions which were postulated at the first meeting, can be backed by scientific and technological advances tested in concrete applications and projects.

3.1 Motivation

Any paradigm shift has to be rooted in substantial arguments backed by scientific evidence that serve as hypotheses for further research which would either support or dismiss the initial arguments. We start by claiming that, in order to break through the current limitations to designing human-like intelligent technology scientists in the area of artificial intelligence need to strike new paths towards capturing the essence of the human mind in its wholeness, also in order to stay scientifically consequent. As strong foundation for the emerging community pioneering these efforts, based on the Forum deliberations (see also the DVD of the [ENF 07]) we provide compelling evidence in this regard by evaluating the status quo while setting up the future research directions. To achieve our goal we need concerted interdisciplinary efforts joining the engineers and psychoanalysts in innovative projects working at the junction of paradigm shift that stretches the boundaries of both disciplines to meet them in creating a new automation science capable to tackle the high complexity that has become pervasive in all areas of automation (as there are security, safety, geriatric care, energy management, and building automation as most important examples) as well as the whole information technology revolution driven world. In approaching complexity it is natural to look at nature – of which we are living proof for the successful ways to intelligent evolution. What follows is our open statement for discussion.

3.2 Challenges in automation

Historically, starting from the trick fountains that measured the time during court hearings in ancient Egypt followed by the "marvels" like Leonardo da Vinci's timepiece, which possessed various clever mechanical processes to measure time in a continuous fashion, automation was concerned with executing processes without the direct influence of humans[61]. The momentum created by the advances in automation led to the industrial revolution – when it got possible to using large amounts of energy for animating devices or controlling processes. With progress in electric generator and propulsion technology, electronic control became perva-

---

[60] This work was supported by the HarrisonMcCain Foundation

[61] Automation mainly aims at taking the monotonic work load from humans, increasing precision, saving energy, and controlling processes that exceed human ability to handle.

sive. What characterizes automation in that time is the mechanical-electrical dichotomy manifested on the physical plane – energy is used to apply the power to either alter something physically or chemically (like move objects or synthesize substances), or to direct or control a process. As with da Vinci's timepieces – within this dichotomy – the flow of energy is directly coupled with the flow of information.

The computer substantially changed these structures and enabled separation of the information flow from the energy flow, Fig. 3.2.1. The current time to be displayed by watches for example at train stations is computed and controlled by a computer. The computer calculates the necessary control signal – the information – for the motor to maneuver the clockhand in finally displaying the current time.

In former times the steam engine was controlled mechanically by a device operating with centrifugal force, which –just like in da Vinci's timepiece – is controlled by mechanical information flows. In contrast, today's control devices, of which the most actual ones, networked embedded control systems, are composed of a nano-scale computing and peripherals that control the actuators via information signals flowing through networks, e. g. so-called eNetworks.

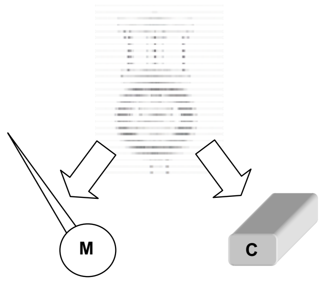

Fig. 3.2.1 Separation of energy flow and information flow bottom left the moving mechanical part (M: Motor), bottom right the respective "intelligence" behind (C: Computer)

This principle seems to be obvious in retrospect, but also evolution needed its time to take this step forward. As an example one can compare the amoeba with a bug as shown in [Die-b 00]. The bug utilizes a dedicated apparatus for controlling itself – the nervous system. If in a next step one compares the information processing architecture of a bug with that of a human, further development in automation is clearly noticeable: data about the processes have to be acquired with evermore, diverse and more precise sensors, and the performance of data processing needs to be steadily increased – which mirrors exactly what happens currently in automation. As the complexity of systems increases with the ubiquity of communication systems and infrastructures, the number of sensors is increasing dramatically, thanks to micro- and nanotechnology and more and more information is being processed and provided (saved) for other tasks. Communications systems not just connect sensors and actuators with their controllers, but also whole networks are connected among each other resulting in systems of systems connected by networks of networks with hybrid characteristics integrated via a unified information communications technologies (ICT) infrastructure uniting all former standalone processes. This enabled among others Computer Integrated Manufacturing (CIM), for 25 years the vision of automation engineers, to become a reality while opening perspectives unthinkable before – to the deployment of systems merging the physical and the virtual into the novel Cyber-Physical Ecosystems that are being applied pervasively in all areas from safety and security to green electricity distribution, vehicular technologies, homecare and building automation [DU 08].

This implies increased complexity. It is usual today that modern buildings have several thousand embedded systems (computers) installed in order to control the process "building" [Die-a 00]. Similarly in modern automobiles hundreds of embedded systems perform their actions, which would be impossible to maintain efficiently or operate safely without the help of automation. Or if the 1-liter-house is to be introduced over a wide area, energy has to be controlled adequately. A parallel with the way humans consume and store energy is a very good example in this respect. Without the nervous system we would not have such an efficient energy household that reacts dynamically to changes and challenges from the outside. And the energy control is just a small part of the process 'human system' necessary to survive. This sheds light on the future directions for automation of which we will focus on two essentials:

Separation of energy and information flow and the requirement to acquire as precise data as possible to allow efficient process control.

However using the human body as source of inspiration is just the tipping of the iceberg. In spite of the enormous investment in developing innovative solutions for intelligent systems, Artificial Intelligence failed to emulate substantial capabilities of natural systems in general, let alone human abilities which thus cannot yet be used by engineering in creating devices that would display similar capabilities. Natural adaptation of the device to the particular human, including the emotional status as well as the instant assessment of situations, anticipating needs or learning from experience are still in the infancy on the technology side. Such capabilities would give automation a new quality of artificial natural-like intelligence that transcends the traditional incremental improvement solution through a disruptive shift in paradigm that originates in the deepest possible understanding of nature in its most intimate information processing principles and solutions. Psychoanalysis offers the knowing and premises for such a radical shift in embedding human-like experience into machines[62].

---

[62] This can be further used as inspiration in the design of large scale adaptive systems by using eNetworks as nervous systems of the ICT controlled 'ecosystems' such as, e.g. 'energy webs' - energy networks capable to tune the energy production and distribution by natural user demand; holistic security ecosystems animated by networked enabled operations bringing together ad-hoc first responders and networked devices/weapons to respond timely to unexpected crises; hazard free transportation

3.3 Model of function vs. model of behavior

Although the discussion whether to use a model of behavior or of function leads to heavy controversies in psychology, it was not articulated at the INDIN/ENF (see the DVD of the [ENF 07]).

While both methods have advantages and disadvantages depending on the area of application from the automation perspective the results are usually mathematical formulas or can be evaluated by practical experiments [Bru 07]. The goal is to design a model which can be used for the development of a device intelligent enough to perform complex tasks, such as for example to be able to recognize a dangerous situation as described in [Tam 03, p. 57] : A small child enters the kitchen and a hot pot is found on the stove. No adult is present.

To design an automated device that would be able to protect the child it is not sufficient to know about the behavior of the process, but it is compulsory to define its functions. Therefore, in the example with the child the important feature of the control device is to recognize and handle the complex relationships in order to anticipate the eventual possible danger. The key aspect here is complexity. It is practically impossible to control such a complex situation by instructing the device what to do at each step – thus the traditional AI methods that would involve for example either the composition of tables with the information, specifying what has to happen in which case or to use rule-based algorithms which define the behavior of reacting to several circumstances. However, such behavior-based descriptions can only be seen as a simplified option for solutions to tasks which doesn't meet the need of complex real-life situations, be they as simple as in the example considered. To approach this, one has to take the next step by analyzing the system and splitting it up into its functional entities. Normally, functionalizing results in an unendingly more complex process which is not feasible to pursue, however it is the only way to come closer to reality[63].

Considering another example leads to the same conclusions: It is hardly possible to describe the behavior of a PC due to its complexity. However, such a description would even not be useful for the endeavor of building another PC. Instead, considering the structure of its functions and their mutual influences – hence, a functional model – would rather do the job.

From this we can conclude that while with a behavioral description it is possible to analyze or verify how a process behaves and under which circumstances, in

---

(automotive networks for aerospace and avionics) and eHealth – homecare and telecare, disaster response and pandemic mitigation [Ulieru 2007].

[63] This was one of the crucial points why Dietrich decided to use psychoanalysis in 1999 and not another scientific direction like for example behavioural psychology [Die-b 00]. Behavioural psychology is able to explain things in an easier fashion, but it has to be questioned, if this kind of modelling is close to reality. On the other hand, the disadvantages of psychoanalysis are its enormous complexity and the fact that with many people it creates oppositions against itself to allow thinking in its way.

order to design one, the *functional model* is required. Once this functional model is in place, it can be evaluated using the behavioral model of the system at hand.

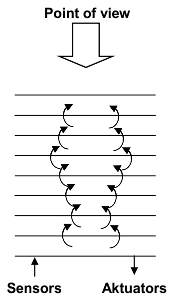

Fig. 3.3.1 Hierarchical model of the brain

In the sequel we will elaborate on why the behavioral model will not lead to success here when tackling the design of complex automation systems. Considering the hierarchical model of [Lur 73] – if it is seen to correspond well with reality, it is important to recognize that the knowledge of the highest (ternary) level – was very limited at his times. Extrapolating his approach, we can define a structure as depicted in Fig. 3.3.1. Additionally, psychoanalysis assumes that the functions in the higher levels are organized as a partial hierarchy which leaves some degree of autonomy to the distributed entities to communicate according to their needs (heterarchy).

It is general knowledge that the human body has numerous control structures, both mentally and physically. According to Fig. 3.3.1 this means that all levels are interconnected and have influences and feedbacks on each other. The control structures within the levels take on stable states which are dependent on their boundary conditions. In this structure, while observing the behavior of the upper level, it is possible that one of the lower levels is altered which leads to significant change in this level. On the other hand the lower control structures manage to keep the system stable – which cannot be considered while observing the upper behavior. One main reason is because the middle layers are not yet known and described, there is just a vague description of their functions [Lur 73].

In this sense it is clear that e.g. statistical analysis of multiple nested control loops, which are additionally mutually connected and non-linear, cannot help in gaining the functional model thereof. This has to be achieved in another way. All this leads to psychoanalysis.

3.4 Psychoanalysis

Artificial Intelligence and psychoanalysis were mentioned together for the first time in [Tur 89; p. 241] and the question was raised, if there should be cooperation between the two fields.

At the WFCS [Die-b 00], Dietrich for the first time presented a model based on consideration from cooperation of those disciplines worked out by his team in Vienna. These were rudimentary considerations, but today we have to question why these considerations came up at exactly this time and if they make sense at all.

Mrs. Turkle [Tur 89] points out the contradictions between the disciplines and analyzes perceptively the root of the clash between the two sciences. In spite of the mutual prejudices, engineers have increasingly embraced the vocabulary of psychologists, without building the necessary foundation that would ensure proper use. In their attempt to fetch ideas from psychology engineers built a theoretical patchwork based on superficial knowledge – given that, in the engineering pursuit it didn't matter it matched the original theoretical psychology (as per [Bre 02, p.44]). One characteristic of the engineering method is to build a mathematical formalism, which in such cases led to useful results (see also [Joh 01], HBM[64]). Given that the task was to produce a solution to some concrete technical problem rather than to build a model of the human mental apparatus – the task proved useful (as confirmed by HBM).

Aside from these solutions, no significant progress on modeling the mental apparatus is notable. We attribute this to the difficulty that engineers would have in catching up with the psychoanalytical education, which in itself takes – according to WAK[65] – longer than the average education of an engineer. Thus, it makes a lot of sense to bring the communities together in a common endeavor when investigating the functions of the human mental apparatus in an interdisciplinary fashion. Even then, it is essential to avoid inherent confusions that may result from the unavoidable clash between the two so different sciences. Therefore it is suggested to start by constructing a unitary model which follows consequently just one school rather than 'mixing-and-matching' insights and concepts from various, also contradicting schools from the humanities as per [Bre 02], where, in spite of the defying colorful mixture of terms and theories without further compatibility investigations put together does in no way account for a functional model of the human mental apparatus. A design procedure following such considerations can potentially help constructing a particular robot, which at most can copy several typical behavioral patterns questionable if we are to consider them 'human-like'. To clarify the bias here, one has to distinguish between what a robot is and can do autonomously and what a human observer projects into what the observer perceives. Moreover, we need to distinguish between human-like appearance (Fig. 3.4.1), human-like behavior (as per [Bre 02]) and human-like *thinking* – which we claim to be a function performed by the human mental apparatus and this definitely cannot be assumed for the two objects presented in Fig. 3.4.1.

---

64 http://www.seas.upenn.edu/~barryg/HBMR.html

65 Wiener Arbeitskreis für Psychoanalyse; http://www.psychoanalyse.org/

Fig. 3.4.1 Are these human-like?

Of course a natural question now is: Why is it psychoanalysis, which is to be considered promising for robot and automation technology? And why is the time ripe for it right now?

In section 3.2 of this article was presented, why statistical models from behavioral psychology cannot provide design policies. Furthermore, humanities – the science disciplines concerned with – among others also – this matter – provide no other convincing and exhaustive functional psychological models than the one originally postulated by Sigmund Freud[66], therefore it is quite straight forward to utilize that. While other mental models provide widely accepted models of human behavior, psychoanalysis is deeply rooted in the functions of human mental control loops as major drive in their actions. Thus, it is the goal of the authors to found a community to build an automatic system capable to encapsulate this powerful mechanism.

With the advent of pervasive information and communication technologies society is undergoing a radical transformation from the command economy to the eNetworked ecosystem characteristic of the eSociety. This pervasiveness of networks linking large, sophisticated knowledge repositories managed by intelligent agents – is becoming more and more a Universal Mind capable to find and answer almost any question. Linked by eNetworks, global enterprises and businesses merge seamlessly into a forever growing open market economy in which dynamic adaptation and seamless evolution are equivalent to survival. The Global Collaborative Ecosystem becomes more and more hybrid, inclusive and capable of almost everything imaginable [UV 08].

Thus, right now there is a high need for strategies to deal with the high degree of complexity that the information technology revolution will bring, as the authors envisioned already some time ago [Die-a 00], [Die-b 00], [Die-c 00], [Uli 07]. In this spirit we want to use recent results obtained on this path to design intelligent automation solutions [Tam 03], [Rus 03], [Roe 07], [Pal 08].

[66] The authors could not find any other functional model during many years of studying literature und talking to experts in the humanities.

3.5 Neuropsychoanalytic inspired model

The clash between psychoanalysts and engineers goes however far beyond their different jargon - into, their different way of thinking and work methods – after all most psychoanalysts work in therapy rather than in theoretical model building, thus for them the natural scientific point of view is of minor relevance. The clash can at most resume in an acceptable compromise, since the perspectives are so opposed that there is little chance to convert any one to the others. For example, engineers often refuse scientific findings, if they don't seem logical to them (e.g. engineers would go as far as to integrate a state machine in bionic models – where it does not belong at all [Gol 07]). On the other hand psychoanalysts often refuse to give up their mechanistic way of thinking (the imagination of mental energy of Freud is still in the daily vocabulary of psychoanalysis, but it contradicts the above mentioned requirement of separating energy flow from information flow).

While the jargon may be dealt with via a learning process, we see the larger obstacle to be in the divergent essence of the two schools of thought. In this respect there are many potential traps and pitfalls, of which we will underline two.

1) While the engineering approach is based on the formulation of a unitary model that is valid for all systems of the kind the model was built for, psychoanalysts instantiate the universal models for each particular individual case. Thus, while for engineering the hypothesis are universal and valid in all cases, for psychoanalysts an explanation or hypothesis can hold true in just one instance, for a particular individual patient. While, for example, to a psychoanalyst the topographical model of Freud (Id, Ego, Super-Ego) is not required to match in all points the unconscious-conscious model, such an inconsistency would be unacceptable to an engineer.
2) The second example: In computer science one makes a distinction between hardware, software, and application[67]. The application behaves according to the underlying software on which it is built (which comprises the algorithms and the programming) and on the hardware on which it runs. Hence, a behavioral description can be formulated for the application – as a function of the particular software and hardware. Looking at the neuron in a similar manner, as a small, particular computer [Ecc 73] (as was proven already in 1976 in [Lan 76]) – the human brain can be regarded as a heavily distributed computer system. With this definition the functional distinction in hardware, software, application and their behavior can be applied also to psychoanalysis. This leads us to conclude that not only is research on the information theoretical aspects of psychoanalytic and neuroscientific concepts done with mechanistic methods, but psychoanalysis also refuses a distinction between the possible behavioral descriptions[68].

---

[67] The application is the utilization view, i.e. the functionality provided to the user.

[68] The authors are aware that up to now there was no necessity, since Freud turned away from neurological observations.

So, although full of promise, we anticipate the path to using psychoanalytical inspired functional models for intelligent systems to be a difficult and stony one.

3.5.1 The new bionic approach

When considering a top down design, psychoanalysis offers two possible models, so-called topographical models. The first model is about unconsciousness and consciousness and the second model is about the Ego, Id and Super-Ego. These are two models describing particular phenomena. They overlap in large areas, but seem to be very hard to unify, as the sustained efforts invested so far in this have proven. To ease the unification task we suggest an engineering approach that starts with simple theoretical constructs, leaving aside in the beginning the model of unconscious-conscious which cannot be easily integrated into the Id-Ego- Super-Ego model thus leaving too many questions open.

As postulated in the respective presentations at the ENF, we suggest that in the first instance the second topographical model will be modularized into its functional entities while clearly defining and specifying the interfaces between them. At ENF (see also the video of the [ENF 07]) raised, but not discussed, was the question if the psychoanalytic model should be integrated – in the sense of Mark Solms – in the neurological model that distinguishes core consciousness and extended consciousness.

The theory of psychoanalysis needs to be experienced by oneself in order to internalize it, which makes it difficult to grasp by the hard core technically trained engineers. However, many areas are based on the Freudian topographical model which is very well functional organized, so one can envision ways towards a common approach by e.g. augmenting the Freudian topographical model (of the mental apparatus) with the neurological model of Alexander Luria [Lur 73] (of the brain) as functional description of the perception system, which results in a holistic model to which computer engineers can relate. Hence, this modular approach can be further subdivided into functions which can be investigated separately while they still interact with the rest of the model in an orchestrated, harmonious manner as stated per [Dam 99, p.154].

Now let us see how one can tackle the concept of consciousness with these basic considerations.

3.5.2 Core Consciousness

The representation field constitutes the main part of the core consciousness module [ENF 07]. The inputs thereof are data (images and scenarios, which last for just moments in time [Dam 01 p. 29]), which are gained via symbolic relations between sensor values. These inputs are associated with other images and scenarios from memory and weighted by the emotions. The resultant generates new evaluations that are further imprinted to the newly stored images and scenarios. Over time everything is matched within core consciousness and therefore generates (emotionally weighted) reactions [Dam 01 p. 29] - there are no other functions performed by the core consciousness module, even time or history does not play a role here (past and future are not taken into account - aka are not perceived)[69]. Only the 'here and now' is of interest. The being is equipped with this 'core' functionality to react efficient and optimal to outer circumstances.

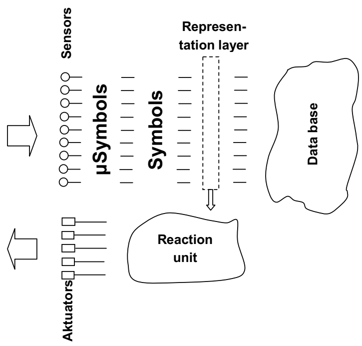

Fig. 3.5.1 Information flow from sensors up to the „data base"

The core consciousness module depicted in Fig. 3.5.1 – does not possess feelings or consciousness in the colloquial sense – thus the reaction unit is not very large. The main part of the core consciousness module consists of the perceptions data base which stores previous seen images and scenarios. It can be assumed that this data base is very small or does not even exist in case of primitive creatures taking into account their relatively small number of sensors. Such creatures are considered to be "hard-wired" or pre-programmed through their genes, although biology still has to confirm this assumption. The networked structures going from the sensors towards the database might rather than extend in fact compress [Bur 07], [Pra 06], [Vel 08], and [Bru 07] while the structures of the outputs can be imagined the opposite way, considering that the number of actuators is much smaller than of sensors.

Thinking of modern robots with capabilities like dancing or riding a bicycle, their intelligence might be rather small compared to primitive creatures. The capabilities of even the most intelligent robot should not be overestimated in spite of the huge sophistication of the involved control algorithms required to perform these functions. A bee might possess more intelligence, but, as mentioned above, this is still an open question for biology.

So, what is the advantage of a core consciousness module compared to the usual control loop algorithms? Core consciousness can be seen as lying above the basic control loops in bodies of creatures like the reflex loop. Just like a control algorithm, core consciousness provides a mechanism that decides how the body should react (in automation such a system is called real-time system). However, unlike in many technical systems, the data is not being analyzed directly. Instead

[69] E.g. a worm does not know about yesterday or tomorrow, it is just aware of the action about to happen now.

perception is essentially based on the comparison of incoming symbolic data which associate inner scenarios and images and directs them to the evaluation system – as per Fig. 3.5.1.

3.5.3 On Embodiment

With these considerations in mind we will attempt to answer a fundamental question: Do we need embodiment when trying to simulate or eventually emulate the mental apparatus? Regarding the mental apparatus to be the evolution's solution as the control device for living creatures, with the ultimate purpose of survival of the individual (body) which is imminently related to the survival of the whole species, it results naturally that *all* capabilities of the mental apparatus aim at fulfilling needs of the body (or the social entity). In this light intelligence modeled on the mental apparatus cannot be thought of without the needs of the respective body. With respect to modeling such a 'mental apparatus' for technical devices, the body can be defined as functional entity and interfaces between 'body' and 'mind' can be specified. Such interfaces and closely related functions can be structured hierarchically [Luria 73].

3.5.4 Extended Consciousness

For the extended consciousness functionality according to [Dam 99, p.195] the principles are basically the same as before, but have to be considered in the more complex context in which the mental apparatus additionally takes past and future into consideration, which makes the *control system* "mental apparatus" a complex, multiple nested control loop containing several underlying control loops[70], as depicted in Fig. 3.5.2. Extended consciousness takes into account previous perceptions and evaluations and *relates them to oneself* in finding feeling-based[71] decisions that further match to previous experiences. Feelings in this definition are evaluations again, but in contrast to the emotions identified in core consciousness, feelings always evaluate in relation to the impact on oneself. To core consciousness, recognition or evaluation is not available (A bee for example, of which can be assumed to possess just core consciousness but no extended consciousness, does not feel pain, it just perceives it.).

---

[70] Fig. 3.5.2 is a very coarse depiction, which can be refined or extended depending on the aspects one wants to enhance it (e.g. with the hormones system of humans).

[71] English language just knows the term *emotional*, although in this case feelings are involved.

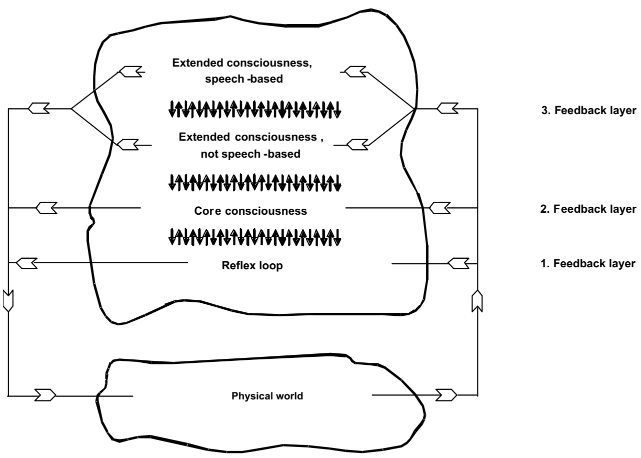

Fig. 3.5.2 control hierarchy

In this top level control loop two representation layers, independent of the one hold by core consciousness, are assumed: Humans have a mental representation of the physical world. This includes both, objects outside the body and the body itself. These two together form the outside world, while the mental representation thereof is (part of) the inner world[72]. Additionally, humans have a representation (imagination) of themselves, the *self*. The self is also part of the inner world. In humans consciousness seems to originate over the relationships between the self and recognized objects from the outer world. Experiences about this relationship create feelings. The major problem with modeling this concept is that the self is no fixed representation, but dependent on various influences [Sol 02, p. 279], like the state of the homeostasis, psychic condition, outer influences, etc. The self is in an ongoing flow and will never arrive to a state again which it has lived through once already.

With this, another hypothesis can be stated: In robotics researchers are inclined in modeling state machines [ABV 07]. For the mental apparatus, this approach has to be rejected, because, in light of the above one can never reach the exactly same state a second time.

The extended consciousness module has even more capabilities. Humans can not just look back in the past, but are also able to plan the future, to think and try out [Dam 99, p.197]. This means that based on experiences creatures with extended consciousness think of alternatives in some kind of movie-like imaginations in order to find most advantageous reaction to a given situation.

3.5.5 Thinking

This trial-action is what is called *thinking*. How this works in detail, is yet unknown. According to Solms [Sol 02, p.209] this process is related to dreams, in

---

[72] The terms *outside* and *inside* world are seen in relation to the mental apparatus. Therefore, the body is outside, too.

which the sensory input is just generated as desired and real sensor values from outside are filtered. This raises the following questions to the computer scientist: do these trial-actions run in parallel, or do they run sequentially? How are they initiated? How terminated? What is the base for decisions? Is the evaluation only through feelings[73]?

It is of the essence that these ‘trial actions’ follow pathways that were learned in the past. This means that all the time current perceptions are overlaid with patterns from childhood – all unconscious. Since human perceived the world is very different with the eyes of children, the adult world is perceived partially very biased. If one wants to fully understand what someone else means, all the interpretations will be rooted in interpretations taken by her own acting-as-if, her own learned pattern, which obviously cannot fully comply with the ones of the person to be understood. Interpretations of the behavior of another person can principally be no more than estimates. We have to accept that everybody lives in her own world. Ultimately, these considerations lead to a definition of the range of functions of language: it shall transport the subjective imaginations of the speaker (including sensations from sensory modalities) to allow the listener to interpret as desired by the speaker. Since this is in principle not always possible – because of the different past and therefore the different allocation of terms – many interpretations remain fragmentary.

What the presented concepts of reflex arcs, core consciousness and extended consciousness of humans (Fig. 3.5.2) reveal to the control and automation engineer is a hierarchical concept of nested feedback loops, in which every layer needs the lower layer, and every higher layer logically serves with an ever higher functionality – and reacts ever slower – following control theory in this respect. Going into more details one will soon recognize that the triple pack in Fig. 3.5.2 needs to be further subdivided. Consequently, each layer with enhanced “intelligent” capabilities for ever more complex processing will react with ever lower speed of response.

3.6 Consequences and Conclusions

The ENF Workshop has ignited an unprecedented dialogue between engineering and neuron-psychoanalysis that not only showcased the enormous progress in brain sciences over the last decades but as well demonstrated that psychoanalysis can serve as a strong foundation for the development of extremely useful engineering models.

In spite of the natural opposition coming from the mainstream engineering school of thought – that is challenged in its core principles by the new school very disruptively – we can already easily dismiss several aberrations attempting to solve the new, complex problems with the old engineering school of thought paradigm - still sucking lots of funding into dead-end research. As a generic example

---

[73] The term *feeling*, which according to Damasio clearly describes an evaluation instance, may not be confused with the colloquial term feeling – which has quite diffuse meanings.

consider Fig. 3.5.1, derived from the work of Solms and Damasio, showing that imaging technology-based brain investigations can only provide analysis of simple tasks involving neurons nearby sensor data, such as e.g. if one decides to go right or left, or to raise a hand, etc - when a mechanism similar to a device driver[74] – in the language of computer engineers – has to be executed by the mental apparatus – which, for such simple 'decisions' can thus be localized in the brain.

However if more complex considerations are undertaken in higher cognitive areas, such as speaking, feelings or thoughts, respectively plans, this involves the whole brain – which was already postulated by Sigmund Freud 100 years ago [Fre 01]. The idea of finding concrete neurons on which thoughts can be read contradicts neuro-psychoanalytical models being as unfeasible as the idea of finding the behavior of a complex system by analyzing the behavior of its parts [Uli 07].

Another aberration that our work dismisses on similar grounds - is the idea that neuron-imaging can help investigating consciousness. Even if it would have the resolution to depict the impulses of single neurons or synapses - imaging can only be of auxiliary help - while often lacking scientific sustainable interpretations especially if not restricted to the lower neurological layers. A trivial example from computer science easily makes our point: it is as if attempting to analyze the functionality of the text processing program WORD (by Microsoft) by scanning myriads of transistors (≈ synapses) from the computer hardware. (And let's take note that a computer is much (!) simpler in its design than the human brain …). Any simple function of WORD – such as inserting a letter – triggers bunches of transistors thus leaving the researcher puzzled in its attempt to explain the complex functionality by looking at the behavior of the elementary components (transistors). The fundamental bias here is that one cannot achieve complex functionality by instructing every elementary element at each step exactly what to do. The attempt to find which transistors account for the respective function is an impossibility. Signals would be overseen in the whirlpool of events. Such a synthesis reveals though that below the application software (WORD in our case) are the operating system, device drivers, etc. Thus to understand how WORD is processed in the computer (aka complex thought process in the brain), one first has to understand the operating system – which luckily we are not required to do beyond the university undergraduate benches. The idea here is that to synthesize behavior one has to scale the system by clustering the components accounting for various functions – which can be done at several levels of resolution [UE 04]. Such considerations - also with implications on philosophy – make it relevant why humans will never be able to recognize reality, because they always think in models via the association of images and scenarios which they already learned. We live in a subjective world, the objective world will always be only accessible to us indirectly through our models

Another point worth mentioning here is the quite popular view in brain research that the brain does not work like a computer – which in fact is not based on

---

[74] A device driver in computer science is a program that allows other parts of the software access to hardware.

any scientific evidence. Even if one is to look at it in the most popular accepted view[75], as a device that processes information [Dud 89] [Hay 88] [Pat 94] – a computer is certainly very similar to the brain!

A brief look at the history of computers further confirms this: Today accepted as the first computer worldwide is the Z3 built by Zuse in 1941 which was a fully automated calculator for binary floating point arithmetic. From an automation perspective Z3 is a device with one single task: to process incoming information and use the results to control processes. Recalling now the process clock from Fig. 3.2.1 we would talk about a device, which provides the mechanical power to operate the clockhand and a computer to control how that has to be done. Based on this preamble, a computer can be defined as a system which processes information. Since ’41, there have been many attempts to enhance the computer’s functionality, starting with analogue calculators and threshold systems to the nowadays attractive quantum computer. However, if we do not award the brain with metaphysical phenomena, it is very hard to find a principal difference in task and function between the brain and a computer.

If we are to truly achieve the breakthrough in computing that will enable the processing of complex tasks involved in the 21st Century problems [BS 08] we will have to break through the current paradigms and redefine computing and its abstractions [Lee 07] in terms that will enable it to conquer the new challenges. The ENF community strongly believes that the psychoanalytical perspective which is convinced to allow a natural scientific top-down description of the functions of the mental apparatus– is one way towards achieving this breakthrough. Thus we are determined to capitalize on the latest achievements in psychoanalysis in building the new computing abstractions for the next generation (fifth generation of Artificial Intelligence) intelligent machines. Such abstractions would – in addition to the future human-friendly technology and environments – be extremely useful in the design of the much needed ‘intent analysis’ systems [Wer 00] capable to capture e.g. the intent of a human – being this information used for increasing comfort, security, safety, or else. Such systems incorporating psychoanalytic insights could potentially also be extremely useful to help in grasping the (religious, cultural, ethnic) differences that generate human conflict thus opening the opportunity to help in mitigating them – thus increasing the chances for peace on our planet.

With this said – we, the ENF community are ready to redefine artificial intelligence, cognitive science and brain research by bringing the insights of neuropsychoanalysis into the engineering world. After all – with every second that we delay this endeavor – our world misses …

Literature

[ABV 07] Argall B., Browning B., Veloso M.: Learning to Select State Machines using Expert Advice on an Autonomous Robot; In Proceedings of ICRA'2007, Rome, Italy, April 2007

[Bre 02] Breazeal, C. L.: Designing Sociable Robots; The MIT Press; 2002

[75] http://de.wikipedia.org/wiki/Computer

[Bru 07] Bruckner D.: Probabilistic Models in Building Automation: Recognizing Scenarios with Statistical Methods, Dissertation Thesis, Vienna University of Technology, 2007

[Bur 07] Burgstaller W.: *Interpretation of Situations in Buildings,* Dissertation Thesis, Vienna University of Technology, 2008

[BS 08] Systems Thinking: Coping with 21st Century Problems, CRC Press, ISBN 978-1-4200-5491-0

[Dam 99] Damasio, A. R.: The Feeling of What Happens: Body, Emotions and the Making of Consciousness; List Verlag (Econ Ullstein List Verlag GmbH); 1999

[Die-a 00] Dietrich, D.; Loy, D.; Schweinzer, H.-J.: LON-Technologie; Verteilte Systeme in der Anwendung; Hüthig Verlag, 2. Auflage, 1999

[Die-b 00] Dietrich, D.: “Evolution potentials for fieldbus systems”, IEEE International Workshop on Factory Communication Systems WFCS 2000; Instituto Superior de Engenharia do Porto, Portugal (2000)

[Die-c 00] Dietrich, D.; Kastner, W.; Sauter T. (ed.): EIB Gebäudebussystem; Hüthig, Heidelberg, 2000

[Die 04] Dietrich, D.; Kastner, W.; Schweinzer, H.: Wahrnehmungsbewusstsein in der Automation - ein neuer bionischer Dekansatz; at, 52. Jahrgang – März 2004, pp. 107-116

[Die 07] G. Pratl, D. Dietrich, G. Hancke, W. Penzhorn: "A New Model for Autonomous, Networked Control Systems"; IEEE Transactions on Industrial Informatics, Volume 1, Issue 3 (2007), S. 21 - 32.

[Dor 01] Dornes, M.: Der kompetente Säugling - Die präverbale Entwicklung des Menschen; Geist und Psyche Fischer; Fischer Taschenbuch Verlag, 2001

[DU 08] Doursat and Ulieru: Design by Emergence: The Challenge of Engineering Large Scale eNetworked Systems, IEEE Transactions on Systems, Man and Cybernetics - Part A, Special Issue on Cyber-Physical Ecosystems (to appear in fall 2008)

[Dud 89] Duden - Informatik; Dudenverlag, Mannheim, Wien, Zürich, 1989

[Ecc 73] Eccles, J. C.: The Understanding of the Brain; McGraw-Hill Book Company, New York; 1973

[Ecc 94] Eccles, J. C.: How the Self Controls its Brain; Springer Verlag Berlin, Heidelberg, 1994

[ENF 07] Conference Proceedings ENF 2007; 1st International Engineering & Neuro-Psdychoanaysis Forum der IEEE/INDIN 2007; Dietrich, Fodor, Pratl, Solms; Vienna; July, 2007
DVD: to buy by Dietmar Dietrich: http:// www.ict.tuwien.ac.at

[Fre 01] Freud, S.: Zur Auffassung der Aphasien; Eine kritische Studie; Herausgegeben von Paul Vogel; Psychologie Fischer; Fischer Taschenbuch Verlag; 2001

[Gol 07] Goldfarb L., Gay D., Golubitsky O., Korkin D., Scrimger I.: What is a structural representation? A proposal for an event-based representational formalism; http://www.cs.unb.ca/~goldfarb/ETSbook/ETS6.pdf; 2007

[Hay 88] Hayes, J. P.: Computer Architecture and Organization (2nd ed.). New York, NY: McGraw-Hill Publishing Company, 1988

[Joh 01] Johns, M.; Silvermann, B. G.: How Emotions and Personality Effect the Utility of Alternative Decisions: A Terrorist Target Selection Case Study; 2001

[Lan 76] Lange, D.: Aufbau zweier Neuronenmodelle mit modernen Bauelementen nach Küpfmüller-Jenik zur Nachbildung gegenseitiger Erregungs- und Hemmvorgänge bei neuronaler Verkopplung über exzitatorische beziehungsweise inhibitorische Synapsen; Diplomarbeit an der TU Karlsruhe; Betreuer W. Adrian; 1976

[Lee 07] Lee, A.E, Computing Foundations and Practice for Cyber-Physical Systems: A Preliminary Report, UC, Berkeley, UCB/EECS-2007-72, May, 2007

[Lur 73] Luria, A. R.: Working Brain: An Introduction to Neuropsychology. London, UK: Penguin Books Ltd., 1973.

[Pal 08] Palensky, B.: *From Neuro-Psychoanalysis to Cognitive and Affective Automation Systems*, Dissertation Thesis, Vienna University of Technology, 2008

[Pat 94] Patterson, D. A.; Hennessy, J. L.: Computer Organization And Design - The Hardware/Software Interface; Morgan Kaufmann Publishers; San Mateo, California, 1994

[Pra 06] Pratl (Zucker) G.: *Processing and Symbolization of Ambient Sensor Data*, Dissertation Thesis, Vienna University of Technology, 2006

[Roe 07] Rösener Ch.: *Adaptive Behavior Arbitration for Mobile Service Robots in Building Automation*, Dissertation Thesis, Vienna University of Technology, 2007

[Rus 03] Russ G.: *Situation Dependent Behavior in Building Automation*, Dissertation Thesis, Vienna University of Technology, 2003

[Sol 03] Solms, M.; Turnbull, O.: The Brain and the Inner World: An Introduction to the Neuroscience of Subjective Experience. New York, NY: Karnac Books Ltd., 2003

[Tam 03] Tamarit C.: *Automation System Perception – First Step towards Perceptive Awareness*, Dissertation Thesis, Vienna University of Technology, 2003

[Tur 89] Turkle, Sh.: Artificial Intelligence and Psychoanalysis: A New Alliance; editor: S. R. Graubard; MIT Press, Cambridge, MA; 1988

[UE 04] Ulieru M. and Este R., "The Holonic Enterprise and Theory of Emergence", *International Journal Cybernetics and Human Knowing* (Imprint Academic), Vol. 11, No. 1, 2004, pp. 79-99.

[Uli 07] Ulieru M. Evolving the 'DNA blueprint' of eNetwork middleware to Control Resilient and Efficient Cyber-Physical Ecosystems, Invited paper at BIONETICS 2007 - 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems, Budapest, Hungary December 10-14, 2007.

[UV 08] Ulieru M. and Verdon J. "IT Revolutions in the Industry: From the Command Economy to the eNetworked Industrial Ecosystem" - *Position Paper at the 1st International Workshop on Industrial Ecosystems*, IEEE International Conference on Industrial Informatics, Daejoen, Korea, July 13-17 2008.

[Vel 08] Velik R.: *A Bionic Model for Human-like Machine Perception*, Dissertation Thesis, Vienna University of Technology, 2008.

[Wer 00] Werther, G., Profiling 'Change Processes' as a Strategic Analysis Tool, Competitive Intelligence Magazine, January-March, 2000