New Approach for Controlling Complex Processes. An Introduction to the 5th Generation of AI

Dietmar Dietrich†, *Senior Member, IEEE* and Gerhard Zucker (né Pratl)†, *Member, IEEE*
†Institute of Computer Technology, Vienna University of Technology, Vienna, Austria
dietrich@ict.tuwien.ac.at, zucker@ict.tuwien.ac.at

Abstract — **In automation mechanical units are more and more replaced by electronic components. They contain sensors, actuators and information computing systems. The goal is to achieve higher process and product quality. That means, for example, to increase the performance by implementing more functionality. Systems become interconnected, which causes increased complexity – something that cannot easily be handled with traditional technological approaches. It is necessary to rethink existing approaches and find new solutions. These could be found in bionic models. Contemporary researchers like Mark Solms and Antonio R. Damasio build upon psychoanalysis, which could be such a new approach for engineers. The theories of psychoanalysis are widely unknown to engineers, still their methods, which originate from neurology, are partly similar to the methods of computer technology. It is merely the theoretical foundation of psychoanalysis, which is complex and at the start hard to grasp for an engineer. Cooperation between engineers and psychoanalysts thus has to overcome this hurdle before being successful. But if we want to conduct our scientific work consequently, it is high time that we get acquainted with psychoanalysis. The following paper wants to show how to install such cooperation and tries to outline the idea of a model that results from already existing collaboration.**

Keywords — **automation, bionic model, control of complex systems, psychoanalysis**

I. Motivation

Antonio R. Damasio and Mark Solms [Dam 99], [Sol 03] are two distinctly different researchers. Damasio is a classically educated neurologist, while Solms is psychologist and psychoanalyst. Today they are the top researchers in brain research. Both come to the conclusion that emotions and feelings, which have to be considered separately, serve a special role in the history of evolution: an individual that is equipped with emotions and feelings is superior to others, because it can contemplate about its past, envision future events and thus is able to plan its actions. This ability requires at least a minimum amount of processing infrastructure – neurons in the biological world – which is why this ability cannot be adjudged to, for example, bugs.

If we want to be able to handle complex control systems that we are facing today in automation – and which has been the dream of artificial intelligence for the last 50 years – then we should learn to understand how the mind works. Especially the higher, more abstract levels are relevant; processing in the lower layers – the neurons – are in this context less interesting, since a technical solution will not ground on a biological base, it will instead use electronic computers.

This will allow us to design technical systems that are able to control complex processes: surveillance systems for airport security to detect dangerous scenarios; efficiently interpreting millions of data from some thousand embedded systems in a building or in a production plant; providing support in a household and especially supporting the elderly in geriatric applications; controlling trains and airplanes. The principle behind all these applications is identical: how can a machine perceive its environment and interpret the perceived information on a high, abstract level and draw reasonable conclusions?

A model that can describe human behavior is not sufficient. A human being is much too complex that we could infer its function directly from its "random" behavior. It will never reach a state that is identical to a state in which it was in the past – there are no two identical states in one life. The question for an appropriate functional model is also not answered in the description of synaptic connections between neural structures, since this description does not explain how a human "functions", i.e. it does not explain the functional modules that it contains. To give an analogy: if we have a description of the behavior of a word processing program, it will not suffice to deduce the underlying functional modules of the program; even worse, if we examine transistor structures, we will not be able to describe the functionality of a word processing program, too much complexity lies between the operation of a silicon chip and a software application. Still an electronic computer system is by far simpler than a biological brain.

If we look back at the development of artificial intelligence[1], we see the great achievements that have been produced. Norbert Wiener was already aware that nature is the source of answers. We have to be ready to accept results from other scientific disciplines.

But we also realize that we have made some big

[1] When we refer to AI we also include cognitive science and see it as a subset of AI. This is only to simplify the text for the reader, we are well aware that the two communities do not perfectly overlap.

1-4244-1543-8/08/$25.00 ©2008 IEEE

mistakes in artificial intelligence and, in our opinion, sometimes left the grounds of science. This should now be corrected by allowing alternative considerations beside the traditional approaches.

II. Bottom-up versus Top-Down Design – Behavioristic Model versus Functional Model

According to [ENF 07], [Pal 08] we can distinguish four generations of artificial intelligence (AI): First, symbolic AI, which works on symbolic coding, manipulation and decoding of information. Symbolic processing methods are today an essential base for AI, which is also shown in the work of [Vel 08].

Statistical AI employs networked structures and their couplings. The central topic is learning, which unfortunately always only leads to one or more optima, but does not go further. Learning from what has been learned can not be modeled – the system is not able to change its learning structures based on what has been learned. This is also understandable from a different viewpoint, since only the ability to permanently store images and scenarios and integrate them into the learning process can yield the desired performance and diversity. We will discuss that a bit further below.

Next we realized that the body with its requests for satisfaction of needs is a necessary requirement for intelligence – this is the discipline of embodied intelligence. Intelligence does not end in itself; it serves the body to subsist efficiently, as Antonio R. Damasio states in [Dam 99]. Seen from the philosophical viewpoint this was great progress, because this step had significant consequences on the discussions of the existence of a soul (Hegels body-mind problem). If we work scientifically, it becomes clear that thinking and body are the same thing, only seen from different viewpoints. This is exactly the statement of modern psychoanalysis and where neurology meets psychoanalysis and merges into neuropsycho-analysis [Sol 03].

Finally, at the end of the 1980s it was understood that the brain has to be more than just a system on which different programs run for different purposes [Pal 08]. Emotions, language and consciousness are properties that are of great relevance for human beings. This is the domain of emotional AI. First there was no separation between emotions and feelings and we simply assumed that these are evaluation mechanisms. Still phenomena like *consciousness* or the sensation of *feeling something* could not be explained.

Analysis of this historical development makes two things clear. First we see that engineers have chosen a bottom-up method for their designs, simply due to lack of knowledge. We put one piece on top of the other to get from neurons (which would be equivalent to hardware in computer models) to consciousness (which is the highest, most abstract level). But we should not do that; we should always work top-down like computer engineers are used to work. Second, we have selected certain structures, functions or behavior from the whole complex of human abilities, and tried to implement these in technical systems. We expected to understand the rest, which we left out, in good time. Especially the last two generations of AI have used psychological knowledge, which was, however, not checked for consistency. This resulted in a patchwork – a mix of different, partly contradicting schools that were all merged together into one model. Such an example is shown in [Bre 02], which violates scientific principles. If psychological results of different schools are used that have not been checked for consistency (or interoperability, to use a more technical term), the results are worthless.

This must have implications for future research, which, in our opinion, leads to the 5th generation of AI. The boundary conditions are clear:

1. Computer engineers must demand top-down design.
2. We need a *uniform* model, which is free of contradictions.
3. Engineers have to cooperate with the scientists, who have always dealt with the brain and the mind.

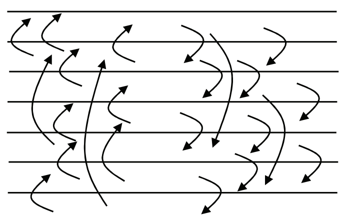

Fig. 1: Layer model of the brain with reference to a computer

Computer engineers successfully use the layering shown in Fig. 1. Such a layer model is currently not available for the mind, but we have to try to apply the layer model anyway. If we work top-down, however, it becomes clear that an abstraction of the neural network (the hardware, to compare it with the computer domain) is not necessary at the beginning; we will explain that a bit further below. We do not know how the brain is "programmed" on the neural level. And even if we knew, it would not be on the right level of modeling; we have to focus on the psyche, and we have to remember that modeling is a multi-layer process (Fig. 1); therefore we cannot map specific behavior of a process (an "application" in the computer domain) directly to functional modules. In doing so we would reduce the complex system "human being" to a primitive system, which contradicts reality: the control unit of a human being does not execute the same behavior twice absolutely identical, since conditions always change. Today's robots (which, in this context, are the same as puppets) have one thing in common: they are not human, but it is the human observer, who projects human-like properties into them. According to [Sol 03] human beings consist physiologically and psychologically of many different control loops who are hierarchical as well, but nevertheless strongly interconnected with each other (Fig. 1). Observing the behavior within *one* experiment can not lead to a complete description of the system. This is the wrong approach, even if it looks stunningly simple. To give an analogy: biologists do not explain flowers by visual features any more, but they look for functional units by considering the flower as a "process". Similarly, trying to describe merely the behavior of a communication protocol under different conditions is a lost cause, since it will never cover the complete abilities of a communication protocol. But it is possible to develop a functional model of a process (where we use the word "process" to refer to both technical and biological occurrences) and understand its behavior based on this model. If we want to build robots or other systems that are intelligent in a more human-like way we have to get away from behavioral description and employ functional development instead. While this is well understood in computer and communications engineering, some psychological schools like behaviorists have a slightly different view of the matter.

III. Engineering and Psychoanalysis

Engineers repeatedly tried to model parts of the brain by referring to neurological and psychological findings, respectively. To understand neurology by analyzing the structure and interconnections of neurons is a task that suits engineers and should be fairly easy due to similar education. But when we study Alexander R. Luria [Lur 73], who describes the foundations of brain structure analysis, we as engineers understand that it is an incredibly hard task to grasp all interrelations that exist in the brain. Looking at the other extreme, psychoanalysis, which originated in neurology, an engineer will soon find that studying the scientific literature of psychoanalysis is futile, since we completely lack the necessary prerequisites – not surprisingly, since it takes almost twice the time to finish psychoanalytic education compared to studying computer engineering at university [WAK]. The illusion of thinking that an engineer can understand psychoanalysis without studying it for many years is therefore out of reason – just like we would not take attempts of a psychoanalyst for serious, who tries to apply Maxwell's equations in his research. Additionally, we have to consider that completing a degree in one field does not make one an expert in the field. What remains is that we have to cooperate in a team of neurologists, psychologists and psychoanalysts. But which of these fields is most promising?

Let us consider Fig. 2 showing different disciplines dealing with brain research. Philosophers essentially process knowledge gained in natural sciences and thus cannot supply information, which is relevant for our field. Human sciences can almost not contribute to modeling artificial intelligence. Information theory does not support us when attempting a bionic approach, because we would first have to know the models for systems as they appear in nature. Artificial intelligence (and cognitive sciences) can not make statements on feelings and consciousness. Neurology is focused on "hardware", and pedagogues are not interested in a model of the psychic apparatus; they are concerned with education and make use of different schools of psychology and psychoanalysis. Psychology only looks at partial aspects or concentrates on behavior; but we explained earlier that behavioral descriptions will not easily lead to a functional model. Furthermore, engineers would have to design this functional model, but they lack the necessary prerequisites. Neuropsychoanalysts aim to build the connection between neurology and psychoanalysis (which eluded Sigmund Freud). Only psychoanalysis remains; it works with clear top-down understanding based on Freud's second topographical. Following the common method in natural sciences Freud's functional model has been modified, discarded, augmented by new hypotheses – in short, it has been heavily discussed and thus matured over time. Scientific conferences show the results and current focus in this area of research [NPSA 07], [IPA 07].

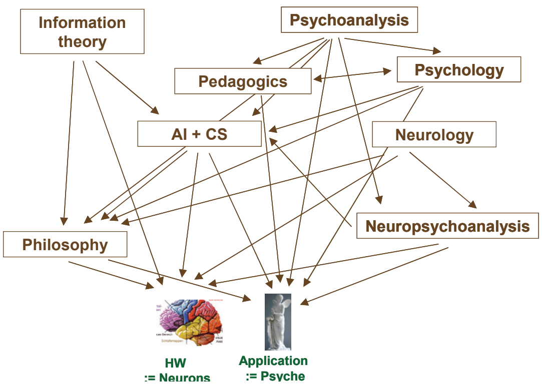

Fig. 2 Disciplines that deal with human brain

This scientific field is often faced with the accusation that it does not work with methods from natural sciences; therefore cooperation is not possible. The following assumptions shall counter this statement:

Assumption 1: We have to decide, whether or not we want to research consciousness. If we do, we have to deal with subjectivity. However unusual it may be for natural science to abandon objectivity, there is no way around it.

Psychoanalysis is accused of not being able to prove anything. To counter this we introduce another assumption:

Assumption 2: Modern sciences are not able to prove everything. Engineers can live well with hypotheses. Just think of latest hypotheses in astronomy or especially the development of the Wankel engine, whose hypothesis had at first not been accredited in science.

Natural sciences therefore have to continue coping with aspects that are not immediately provable and it has to deal with subjectivity in near future. To work scientifically we first have to consider the state-of-the-art, which means to consider *all* scientifically relevant works which have yielded new results in the field. If we deal with *feelings* – which is done by modern AI – we consequently have to sit down with psychoanalysts and discuss how to translate their model into the engineering domain so that we are able to simulate or even emulate it. Scientists in engineering cannot afford to ignore other disciplines.

IV. Flow of Energy and Information

A central question for our research is whether we have to look at the brain as a whole or whether we can understand the mental apparatus on its own, isolated from the brain (i.e. the biological organ). According to Fig. 3, which represents the abstraction of different functions, an engineer understands that hardware has to be seen separately from software; still we are well aware that both are the same matter, and the different functions (of the same matter) interact with and depend on each other.

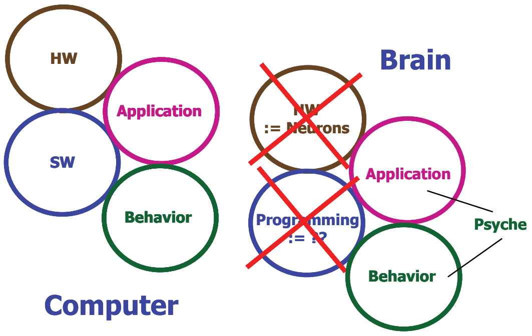

Fig. 3: Comparison of computer and brain

Another important consideration is that information flow and energy flow can be separated from each other (Fig. 4); although this is well understood in engineering today, it was simply not possible in former times[2]. The clock which was designed by Leonardo da Vinci had such genuine mechanics that it was able to display time. The flow of energy was mastered in a way that it could be used to contain the information *time*. This has significantly changed when electronics, especially computers and fieldbuses, were introduced. They represent new media for information processing, while energy, which is necessary to interact physically, is supplied separately. A good example is the Airbus A320, which was the first airplane to abandon mechanical links to control the plane. Computers were responsible for information processing, fieldbuses transported information; the energy to move the flats (actuators with integrated sensors) and the stick (sensors with integrated actuators) was conducted directly to the actuators.

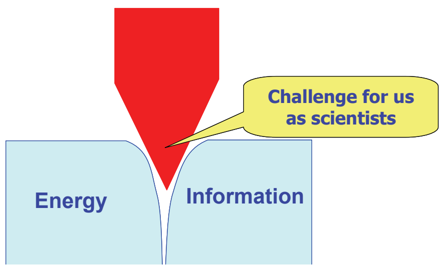

Fig. 4: The flow of energy and information can be separated

This principle is today found in all areas of automation, even in most electrical devices. Separation of information and energy changed technology dramatically. Drive-by-wire is another example that shows how to get a better model of reality. To achieve this, more sensors with more diversity are used; thus the number of sensors in a car has increased dramatically over the last years without reaching saturation; on the contrary, it seems that the process has just started.

Another example is well comprehensible for anyone, who has ever tinkered with radio: software-defined radio. Unlike traditional radio communication systems, which consist of different electrical components like filters and oscillating circuits, it converts analog signals as soon as possible into digital signals and does all necessary processing in the digital domain. The advantages of digital processing are less noise, better precision and much more possibilities for filtering and amplifying. Only shortly before outputting the signal to the loudspeakers, it is converted from digital to analog again. In other words, we use sensors to sample data, which is processed in an abstract fashion, purely mathematical, and no longer physical (or even mechanical) like in Leonardo da Vinci's clock.

The increasing amount of sensors in automation on the other hand creates the problem that more data is created and needs to be processed; it becomes necessary to extract essential information from diverse sensory information. This again raises the request for "higher intelligence" to be able to extract meaningful information.

V. Essential Scientific Finding

Starting from the top-down model we create a functional model of the mental apparatus; a coarse description first separates three modules (in the first step we disregard the model of conscious and unconscious): Freud's second topographical model, the model of perception and the model of memory (Fig. 5).

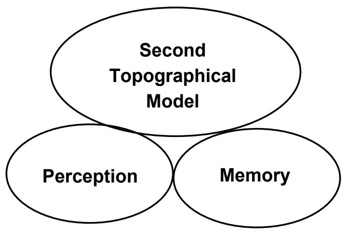

Fig. 5: Modules of the first bionic approach based on the psychoanalytic model

Sigmund Freud disregarded areas of simple perception, which is the part that is not contained in the second topographical model, because he could not access it. The same applies for storage of information. But to create a working system based on the psychoanalytical model we need both modules. That is where we need other scientific disciplines to provide solutions. More detailed explanations are covered in [Die-b 00], [Die 04], [Pra 07] and we only cover one aspect here for better understanding of the following; this aspect is essential and should help to clarify many open questions.

[2] This is also a challenge for psychoanalysis. Sigmund Freud did not manage to remove mechanistic thinking from his model. Analogies like drives producing *energy* are still in use, although they are wrong, since they imply incorrect assumptions.

An important part of perception is to capture images and scenarios. The brain and its sensors do not work like a camera; we do not see the world as separated pixels (we also do not "recognize" separate molecules). Fig. 6 shows an artist on the left. Our sensors (eyes and ears, for example) perceive characteristic features: shapes, areas, phonemes. These are associated in our "database" of existing images, sounds and smells. This again means that the biggest part of what we see, hear or smell is associated with existing memories to create the virtual images shown on the right side in Fig. 6. What we believe to see or hear is merely a construction of manipulated data that we receive from our "database" and only a small part of it is data which is fed in from the real, outside world. We have to separate *images*, which are static combinations of the different sensory perception channels (including olfactory, acoustic and others) and *scenarios*, which are sequences over a short period of time and can be combined from images or be developed independently.

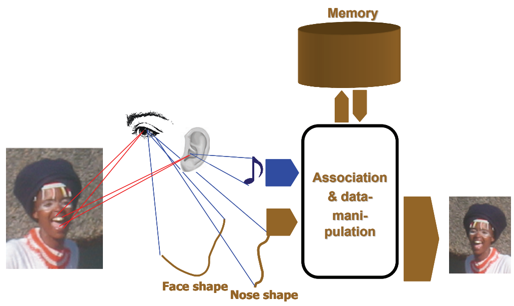

Fig. 6: Seeing the real, outer world [Pra 07]

Scenarios have a limited runtime of a few seconds [Die 08]. Sequences of events that cover a longer period can also be stored and processed, (but this is the task of consciousness) and they are called *acts*. They happen "above" perception, in a system that will now briefly be discussed, before we introduce the bionic model that has been developed by the ARS team to which the authors belong [ARS].

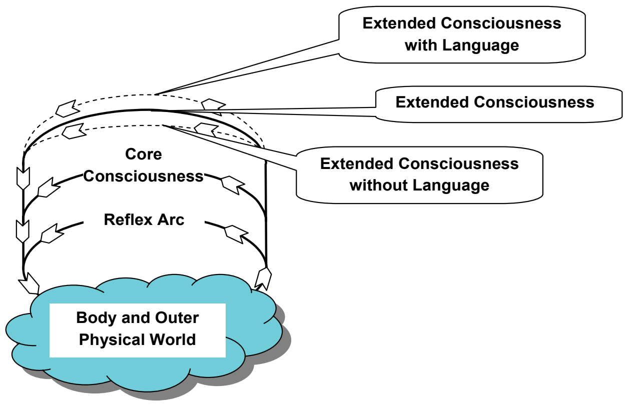

Fig. 7: Hierarchy of feedback loops according to [Dam 99]

Another important hypothesis from brain research, which has commonly been accepted [ENF 07] is the theory of feedback loops of Antonio R. Damasio [Dam 99]. He distinguishes three layers of feedback loops in the nervous system. The lowest layer consists of reflex arcs or direct feedbacks by few neurons like in the spinal cord to control extremities, respectively.

These control loops are overlaid by another control mechanism – core consciousness. It can be assumed that every animal with certain complexity of its nervous system possesses core consciousness. The name itself is, in the opinion of the authors, not very appropriate, since it actually is not really consciousness, but only the core – merely a prerequisite for consciousness. Its task is to see the body as a whole and maintain homeostasis in the body. Core consciousness already uses complex mechanisms of perception that have been described above. Incoming data triggers emotions that lead to certain reactions of the body. There is a huge diversity in reactions available, because it is possible to store and process images and scenarios (i.e. reactions are not based on single sensory inputs that occur right in the moment, but rather on processed and associated information from the present and the near past). Intelligence on this level requires a certain amount of neurons and synaptic connections, which is a necessary prerequisite for the amount of images and scenarios that can be stored. According to Damasio animals, which possess only core consciousness, lack the ability to plan, because they are not able to look back into the past. This requires extended consciousness, which integrates more functions that are well described in [Sol 03]. This is, amongst other things, a layer of representation of one's self and of the inner and outer physical world. Having such representations allows separating oneself from the outer world – an issue that will be covered in the next chapter.

Extended consciousness allows to consciously store images and scenarios into acts, if they are *felt* by the individual. Therefore Damasio strictly separates emotions from feelings. Feelings are properties of extended consciousness. This refers to the ability to consciously look into the past – and also into the future. It allows planning or – as it is phrased in common language – one can think. Based on this Solms explains how, in his opinion, dreams could work and why they are an essential function of the mental apparatus [Sol 03].

VI. New Bionic Model

Based on these hypothetical models it is now possible to create a unified bionic model of the mental apparatus. According to Luria [Lur 73], whose model is still fully valid today [Bur 07], [Pra 06], [Vel 08] the mental apparatus can be divided into three layers: the projection layer, which processes and condenses (symbolizes) sensory information. On top of this layer is the association field, which manages images and scenarios; and the next layer is the tertiary cortex, which is the biggest part of the human brain.

Now we can put together the unified model as follows (Fig. 8): Data from the inner and outer physical world flows through the sensors and is condensed to symbols like *hot*, *oven*, *need for food*, and so on. They lead to images and scenarios, which trigger emotions accompanied by bodily reactions like release of hormones.

The next higher layer creates acts, which can be perceived by the mental apparatus; the mental apparatus is able to connect these acts with itself – and is able to create feelings. Feelings are therefore an evaluation of a situation that is perceived by the mental apparatus.

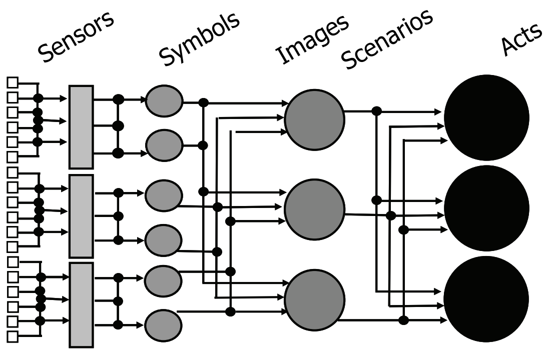

Fig. 8: Bionic Model of the Viennese ARS-team[3]

VII. CONCLUSION

Artificial intelligence had a hard time in the past to make true what was promised. Too many predictions of what will be possible within the next decade turned out to be unachievable. While AI has strongly contributed to both the scientific world and our everyday life, there are still many undelivered promises and it seems that traditional attempts to increase machine intelligence are futile. We have tried to sketch another approach: use existing knowledge from the scientific discipline of psychoanalysis. Although the methods and theories of psychoanalysis are far from easy to understand for an engineer, we do think that it is worth the effort, since we have found it to be the science with the best functional and most consistent description of the human mind. Once we have overcome the first hurdles of common understanding between engineers and psychoanalysts, we can start to integrate their theories into our technical way of modeling. This process has already started in Vienna and has yielded very promising results.

If we are able to create a functional model of the human mind that is described in technical terms and uses only methods and mechanisms that can finally be implemented in a technical system, we have achieved two significant benefits: on the one side engineers are able to build systems, which have the potential to gain human-like abilities. This will only be possible with support from other communities, since there are many partial problems that need to be solved, from machine vision to sensor fusion and mechanics, to name just a few. On the other side, psychoanalysts gain a lot in the process of creating a technical model. Many questions have arisen so far (and will arise in future) only because engineers tried to understand psychological theories with the question "How can we build it?" in mind. To give an example, the question how body and mind interact has been discussed from a completely different viewpoint, because engineers cannot build a body that is made of flesh and blood just to make technical systems human-like. Still we know that a body is the most essential requirement for intelligence. So what is it that we need from the body, which are the core properties of a body that have to be present in a technical system as well?

Research in this field is young and will hopefully continue with lots of work that needs to be done and lots of lively discussions between the different scientific domains. As a result we will be able to better understand the human mind, to create better technical systems and to bring research areas of engineers and psychoanalysts closer together.

REFERENCES

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

[Bur 07] Burgstaller W.: Interpretation of Situations in Buildings, Dissertation Thesis, Vienna University of Technology, 2007

[Dam 99] Damasio A., The Feeling of What Happens. New York, Harcourt Brace, 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 04] Dietrich, D.; Kastner, W.; Schweinzer, H.: Wahrnehmungsbewusstsein in der Automation - ein neuer bionischer Denkansatz; at, 52. Jahrgang – März 2004, pp. 107-116

[Die 08] Dietrich, D., Zucker, G., "Neue Wege der kognitiven Wissenschaft im Bereich der Automation", e&i, to be published in May 2008

[ENF 07] Conference Proceedings ENF 2007, 1st International Engineering & Neuro-Psdychoanaysis Forum der IEEE/INDIN 2007; Dietrich, Fodor, Pratl, Solms; Vienna; July, 2007

[Lur 73] Luria, A. R.: The 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

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

[Pra 06] Pratl, G.: Processing and Symbolization of Ambient Sensor Data, PhD Thesis, Vienna University of Technology, 2006

[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

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

INTERNET LINKS

[ARS] The ARS Project, http.//ars.ict.tuwien.ac.at

[NPSA 07] Congress on Neuro-Psychoanalytic Perspectives on Depression - 19-22 July 2007, Vienna, http://www.neuro-psa.org.uk/npsa/index.php?module=pagemaster&PAGE_user_op=view_page& PAGE_id=48

[IPA 07] 45th IPA Congress, International Psychoanalytical Association, 25-28 July 2007, Berlin, http://www.ipa.org.uk

[WAK] Wiener Arbeitskreis für Psychoanalyse; http://www.psychoanalyse.org/

[3] ARS is short for Artificial Recognition System