Psychoanalytical Model for Automation and Robotics

Dietmar Dietrich, Dietmar Bruckner, Gerhard Zucker, Brit Müller, Anna Tmej

Institute of Computer Technology
Vienna University of Technology
Gusshausstrasse 27-29, A-1040 Vienna, Austria
{dietrich, bruckner, zucker, mueller, tmej}@ict.tuwien.ac.at

Abstract

Research in automation focuses on systems which are capable of solving very complex tasks and problems. Artificial Intelligence and especially Cognitive Science have brought remarkable successes; however, in some areas the boarders of feasibility and further extension are reached. Compared to human intelligence the range of capabilities of the solutions is still modest. In the following we will argue why we see the necessity to introduce a novel approach for creating models, which possibilities and tools computer engineering can offer, why a psychoanalytical template is considered meaningful, and which open problems could be tackled or even broken through with this approach, respectively. The article is based on comprehensive research results in the course of several research projects including a European one. Involved persons originate from a number of research institutions in Austria, South Africa, and Canada.

1. Introduction

In the early days of automation, when the first control components were connected electrically – when first data was transferred over wires – nobody was thinking about protocols. The term was interfaces, and the specifications involved mainly the physical layer. A well known and widespread (mainly in the US today) example is the 20 mA interface. Afterwards, bus systems have been introduced, for which protocol systems in tow layers have been defined. Later on, the ISO/OSI model including 7 abstract layers for defining a communication system has been standardized. Finally, layers for profiles, interoperability, and layers that bring together different fieldbus systems and virtual driver functions finalized the layered architecture of today's communication systems [1].

This development is motivated in the fact that ever increasing numbers of processes get interconnected and should communicate in an automated fashion. Contemporary buildings are already equipped with up to 200.000 embedded nodes [2]. In doing so costs rise for design, integration, and maintenance. For this purpose it is of urgent importance to develop intelligent solutions which focus on identifying process relations, analyzing scenarios, parameterizing of nodes, etc. But building automation as example is only one area out of many. What could be also mentioned is robots involved in complex processes, or the simple example: a child is in the kitchen, a hot water pot is one the stove and no adult is nearby. Why aren't there already "simple" scenario recognition units which scream alarm? Why aren't there already systems which observe airports in order to identify left luggage [3], [4] and analyze if they could be dangerous – as is the goal of a European project [4] which is a side project of the one we present here.

One quickly hits the wall of feasibility if machines should analyze situations or even take decisions. These walls have to be overstepped. New approaches must be sought. In the following, a bionic approach will be presented which could potentially introduce the fifth generation of Artificial Intelligence – based on the ideas of [5].

2. Intelligence of Bionic Systems

Before presenting the developed model it is necessary to explain the motivations why this development was essential. In [6] these motivations could not as yet be formulated in detail.

2.1 Hierarchical Model Conception

In [7] were the various approaches of Artificial Intelligence (AI) investigated. Four generations have been identified since the foundation of AI 60 about years ago: symbolic AI, statistical AI, behavior-based AI, and recently, emotional AI. The article very well describes that researchers started to build and understand functions of the human mental apparatus starting with neurons in a bottom-up design methodology, or even with data-driven statistical analysis. Very much emphasis has been put in a behavior-based design methodology, meaning that devices like robots have been built that behave similar to (move like) humans or animals. The term intelligence itself however has never been precisely defined – even not utilized besides in the title AI.

However, if engineers are to design machines that behave "intelligent" as animals or humans, it has to be dealt with. So, what is intelligence? Is there a general definition? This question can only be seriously answered, if we are able to understand the mental apparatus. It is not enough to know about the behavior, but mandatory to have a unitary and comprehensive functional model of it (figure 1).

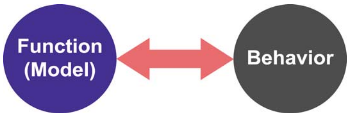

Figure 1: *Differentiation between function and behavior (analytical and behavior-based approach)*

For the later on presented model development a number of boundaries has to be defined. On the one hand appropriate modeling principles and methods from computer engineering, communications engineering, and automation are to be used. Abstract layered models are used there, which are developed in a top-down fashion like for example the ISO/OSI model in communications or the Y-model by Gajski-Walter for digital chip-design. On the other hand, for comparing the platforms, we utilize the definition of a computer, after which it is a data manipulating, storing, and transferring device[1].

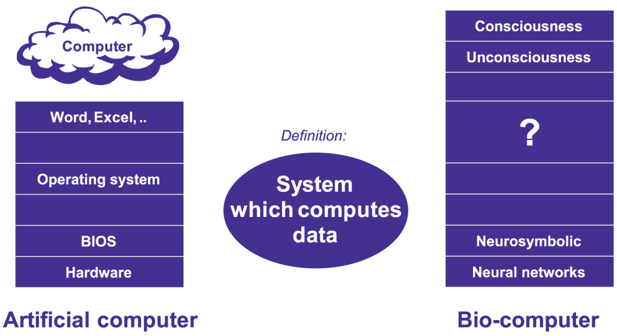

Figure 2: *Possible abstraction layers of a computer (to the left the artificial device, to the right a biological model conception)*

Using that we can base our considerations like in the depiction of figure 2, where there are many abstract layers defined in an artificial computer (from hardware to application software) in the left part – which can be also imagined in the biological model that is shown in the right part, where many layers are not defined yet.

Another evidence for the functional layered architecture[2] of the human brain is explained later and comes from the findings of the Russian neurologist and psychiatrist, the author of [7].

Figure 2 expresses several important issues. On the one hand there is the monist conviction – also along with the natural science world view – that the brain is a control system based on laws of physics, with no principally unexplainable mechanisms[3], no mystics. Additionally, several boundaries can be deducted from the model, which have been partly neglected in previous AI. The next seven sections will go into details about them.

2.2 Statistical Methods

A large number of feedback loops of all kind are acting in humans (figure 4), both physiologically and mentally (hence, in all layers of figure 2). If one now assumes that the behavior of the feedback loop depends on its parameters – if it has more smooth output like in the top right part of figure 3 or more instable appearance as in the diagram beneath – it can be easily imagined that specific propositions about particular feedback loops in such an arrangement can only be stated, if there exists a concrete model conception of the whole process below.

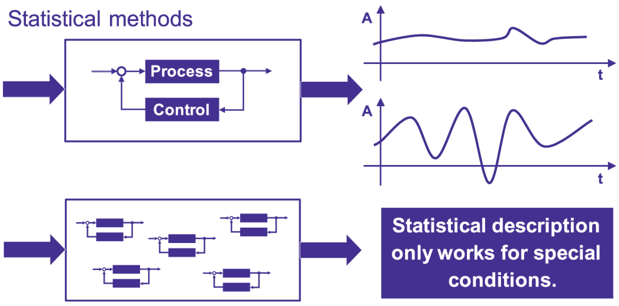

Figure 3: *Feedback loop difficulty*

In this way, two completely different errors in different layers can result in the same behavior. Or, a substantial error in one layer can be compensated by actions in other layers.

So, if the behavior of the whole process is under investigation, then, phenomena observed from the outside can only give indications about feedback loops inside, if a model of the whole process is at hand, therefore an accurate description of all feedback loops must be given.

For this purpose a further small example. The left part of figure 4 is used. If there is an instability or error in the operation system, the observer who sees the application software can state only then specific propositions about the error, if the correlations are known.

Applying this to humans: if human behavior is modeled statistically, but without knowledge about the internal model (different layers), virtually no useful statements about the cause of the error can be given.

[1] It was deliberately refrained from using the more narrow definition which includes the term algorithm, since this one does not reflect some bionic considerations, and we do not see the necessity for this restriction.

[2] The authors are convinced that his model cannot be adopted directly but needs modifications. Justification therefore is provided later.

[3] Even if there are many open issues by today. Maybe they remain open for still a long time …

This is because higher layers within the model can modify or compensate the behavior or even initiate opposite behavior to work against the behavior of the lower layer.

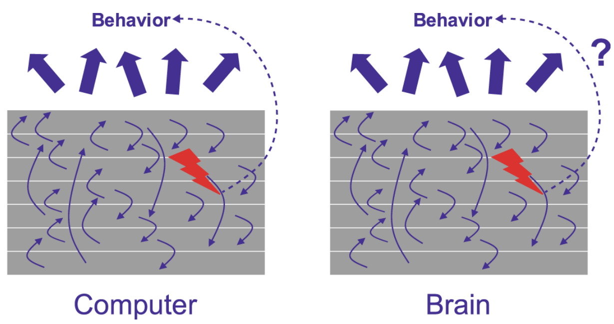

Figure 4: *Effect of layers: different behavior of the whole process*

This description should elaborate why computer engineers (and chip designers) do not use statistical methods of behavior for synthesis. They need an explicit model description[4]. However, AI often uses behavior-based methods, based on statistical analysis for synthesis as can be seen in the example of [8] – a principle, which has to be questioned.

2.3 Definition of Intelligence

The term intelligence has to be used carefully. Some two or three decades ago there have been severe discussions about which microprocessor / computer is the more intelligent – the more powerful – one. However, it was understand soon that this is the wrong kind of question because the computing power cannot be defined in general but only application specific. This means, if a hierarchical model as in figure 2 is assumed, such kind of question can only be applied to a single layer – or a single or several applications have to be defined which are performed by the computer.

So, to define intelligence in a bionic sense (respectively information theory), one has to be aware, of which layer he is talking about. For the synthesis of a model a computer designer needs a conception for every single abstract layer.

2.4 Choice of the Right Model

The huge impact from the choice of the model can be depicted with the following example.

In early medieval times it was assumed the earth is the center of the universe. The description of Mars' orbit (figure 5) was therefore mathematically problematic and possible only approximately with the help of epicycles (ptolimeic world view).

The breakthrough came in 1473 as Copernicus saw the sun as the center – so all orbits could be easily modeled as ellipses.

[4] Statistical methods are however a very important tool if the task is investigation of erroneous behavior.

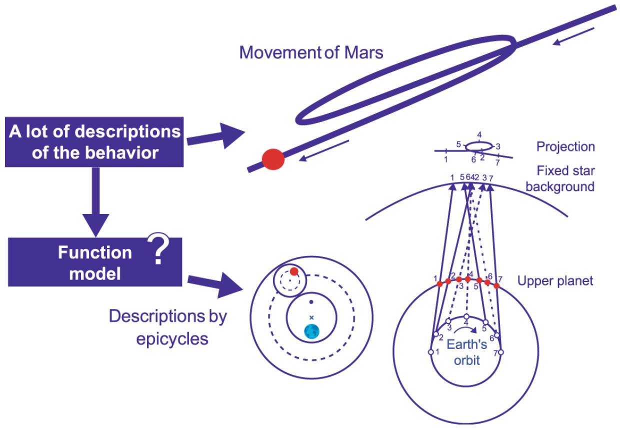

Figure 5: *The earth as the center of the universe*

Transferred to observed human behavior this means that apparently obvious observations and descriptions by an erroneous model tend to give very wrong conclusions. For synthesis of a bionically-inspired model one needs the right model.

2.5 Top-down Methodology

For chip design or computer design one has to distinguish between behavior and function (see figure 1). This means that if a particular behavior is sought, one needs a model of the device under development which can then be iteratively enhanced to reach the desired behavior[5].

Furthermore it is mandatory to use the top-down methodology for the model development, because otherwise the optimal circuit design cannot be found. If, for example, two different behavioral models would be used as basis like in [8], it is obvious that different models will be developed of inconsistencies arise. Here the authors understand top-down as being the design process that starts on the top most abstract functional level – the main function, from which it is divided into sub functions and modules in lower hierarchical abstract layers until a specific description is reached which can be synthesized[6].

2.6 A Unitary Model

Reading about the abundant amount of psychological theories in [8], an engineer has to become sharp-eared. In this book many psychological studies are taken and many psychological publications are cited without having checked their compatibility (interoperability). At least, the authors did not report about this issue, which however is essential, because interoperability has to be guaranteed in the area of interdisciplinary science when taking other theories as template. The various different

[5] All the other way round only experts who have the according knowledge are able to develop such a model – which means it cannot be engineers or physicists.

[6] Being chip designers this methodology has become natural, however, we noticed [9] that other scientific areas have other methodologies and even understand the term of top-down partly totally different.

psychological schools are based on varying premises and this approach to theory-making cannot be compared to natural science practice where it can be seen as major goal to eliminate inconsistencies. In the psychological area it is very hard or even impossible to formally proof a theory, therefore one evaded to more empirical procedures. However, if the goal is to compile the mental apparatus in the bionic sense with methods from computer engineering, one cannot accept inconsistencies, but needs to seek for solutions or for particular application cases.

If one uses results or statements from different psychological schools which are not interoperable, the resulting model conception will be a patchwork, which can seriously not be seen as unitary model of the human mental apparatus.

2.7 Differentiation Between Function, Behavior, and Projection

It was already distinguished between behavior and function. However, having developments like the CB2 [W1] in mind, one absolutely needs an additional differentiation and definition: projection. Looking at the functions of the human mental apparatus, a doll has nothing in common with a human – it simply lacks the psyche to generate any kind of behavior. What makes people call it human-like is its appearance, so people project human behavior into the doll. The same thing happens with robots like the CB2. Even if a number of actuators affect its facial expressions remarkably, there is no emotion of the machine behind. It does not smile because it is amused – it simply lacks the psyche to generate any kind of human-like behavior. It is only the observers that project human-like emotions and behavior into the robot.

CB2 does not show feelings, human beings project feelings into this machine.

2.8 Indispensible Interdisciplinarity

Having a look an international look at working groups on the issue of Artificial General Intelligence, which kind of researchers are involved it can be noted that psychoanalysts or neuropsychoanalysts are actually not present. It can also be noted that there have been some suggestions to bring together psychoanalysis and engineering, however the first investigations in this direction were presented in [5], [9] - [13].

Now there exists the principle in science that all relevant scientific results worldwide have to be included into scientific work. Why has this been not considered, even if psychoanalysis as scientific discipline is concerned with the human mental apparatus for more than 100 years now?

To understand this, another hint is necessary: the education of an engineer in Europe lasts for about 6 years (from going to university), a psychoanalysts needs about 9 years [W2], [W3].

In AI it has become common practice that engineers interpret and utilize psychological literature. On the other hand an engineer would be quite astonished if a psychologist or even psychoanalyst came to the idea to interpret cutting edge technical research articles. The conclusion is clearly that in such kind of working groups interdisciplinary work has to be more emphasized, alone to reduce potential errors[14] – [16].

3. Why Neuropsychoanalysis

On the basis of the above mentioned requirements (top-down methodology, unitary and comprehensive model) an additional premise can be defined. If one accepts that the brain is an information processing system that controls the "process" human, and if one additionally accepts that software is the same as hardware - however in another form of description[7] – one has to conclude that neuronal networks do exactly the same thing, they process information. Following information theory an artificial computer does the same thing, however unfortunately with less performance than the human brain. However, if for the first step of the development only a rudimentary part is to be adopted for the bionic approach, our technically very limited contemporary computers can be used. Hence, for developing the mental apparatus it is not necessary to have neuronal networks, because they represent only the lowest layer (in figure 2) and can therefore be substituted by other hardware, e.g. contemporary electronic circuitry[8].

Since we do not know yet, how the "software" of the brain is to be described, we can only use a behavioral description, or a functionally model – which can be seen as the imagination we have of the human psyche. The analytical approach for engineers has to be always the one with the functional model.

Now, analyzing the various psychological and therapeutic schools (in Vienna, Austria there are already 17 registered [17]), what is left as archetype are only the psychoanalytic schools, since those fulfill the above depicted requirements [18]. Psychoanalytic thinking is based on monistic, natural science considerations following the top-down approach. Additionally, psychoanalysis tries to develop a unitary picture of their models. The fact that the first and second Freudian topological models are not entirely compatible is a scientific challenge, which has to be solved in the future. Further inconsistencies in psychoanalytical theories – which are held against them – have to be seen in the same way. Inconsistencies alone cannot be seen as disqualification. Real problems for engineers however originate from the methods of description, that is why there is still a lot of work to be

[7] Software could in theory be made visible observing the flow of electrons – which would however be very expensive.

[8] This should not be a statement about efficiency.

accomplished for convergence of engineering and psychoanalysis.

So, why neuropsychoanalysis, not only psychoanalysis? Psychoanalysis is concerned with conscious and unconscious mental content, which represent the upper layers in figure 2. For the lower layers (compare [19]), the neuronal networks and the neuro-symbolic networks, neuronal models or models like the one from [20] have to be presented or adopted. This is exactly what the international neuropsychoanalysis society [W4] according to its founder Mark Solms in [20] sees as their proposition. They assume that the gap between neurology and psychoanalysis has to be closed in order to finally acquire a unitary model.

Following these considerations in addition to the once presented in [20] and [21] and incorporating the work of [12] and [13], we have been able to design a functional model, which will be presented in the next section. The basis for the lower levels of the abstract model can be seen in [6], [7] the psychoanalytical model after [20] assumes a very distributed system.

4. The Psychoanalytical Model

The new functional model uses the psychoanalytical model of the human mental apparatus to describe motivation, wish-generation, decision-taking, planning and execution. The main task of the mental apparatus, from a psychoanalytical view, is to synthesize and mediate between three very different demands (figure 6): the drive demand (E3), based on bodily-physiological requirements (E1), the reality demand, consisting of internalized knowledge about facts of outer reality, its possibilities and limitations (E9) on the one hand, and the subjective consequences developed from perception of the outer world (E14) on the other, and the demands of the Super-Ego (E7, E22), which, quasi representing morality, is composed of socio-culturally founded rules and the threats of consequences if these rules are broken.

The drive demand takes the following course: A physiological, hormonal, etc. imbalance (E1) triggers a drive tension (E3) via neuro-symbolization (E2). In a next step, life-sustaining drives are merged with aggressive drive tendencies (E4). The resulting drive content in the form of thing presentations is then connected to and thus rated by an affect of more or less unpleasure (E5). The (still unconscious) representation of the drive content is thus formed by affect and thing presentation.

This drive content, i.e. affect and thing presentation, are transported (I1.5) to the defense mechanisms (E6). Here, the moral rules and threats stemming from the Super-Ego (E7), the internalized reality demands (E9) and the strength of the drive demand are weighed up to in the end decide if and in what form affect and presentation can be handed over to further (and potentially conscious) processing. .

Affects and presentations which are due to Super-Ego or reality-demands not allowed to become conscious or even pre-conscious, are pushed to a container for repressed mental contents (E15).

As yet, all mental processing has taken place within the psychic instance called Id. Thus, all mental content has been subject to the primary process, meaning that the processing is entirely unconscious and does not adhere to any logical rules. Drive contents that have been able to gain access to the Ego and therefore to possibly conscious, i.e. preconscious processing, are now converted to the secondary process (E8). This means that the thing presentations are connected with the according word presentations. Thus, presentations can be ordered and assessed logically, e.g. according to temporal and spatial conceptions.

The reality demand, again, takes its origin in the sensor data of the outer world, consisting of the environment (E10), on the one hand, and the body (E12), on the other. These raw sensor data are transformed into neuro-symbols (E11, E13) and are then unconsciously perceived in the form of thing presentations. These preliminary-perceptional thing presentations activate repressed (unconscious) mental content and together also activate consciously accessible memories (E16). In a next step, the raw perception data and connected repressed content as well as memories are merged to form a perceptional representation (E17). Just like the drive contents, the representations of the outer world produce and are connected to affects (E18). This means that perception does not work with any kind of copy of the outer world, but a subjectively and individually assessed and associated representation thereof. Next, perception content is subject to the workings of the defense mechanisms (E19) in a very similar way as has been described for the drive contents. Again, perception content that pass the defense mechanisms and therefore gain access to the Ego, are transformed to the secondary process (E21).

The affects related to the drive contents and to the perception contents can now be pre-consciously or consciously perceived (E20), which has a large impact on later decision-making and planning (see below).

The drive contents as well as the perception contents are transported (I1.7, I2.11, I5.5) to three further processing entities: the pre-conscious Super-Ego (E22), the decision unit (E26), and the attentive outer perception (E23). The difference between the pre-conscious Super-Ego and the unconscious one, is that the rules of the pre-conscious Super-Ego are more easily accessible and thus closer to consciousness. It resembles the more familiar term of “conscience” in the way it takes an influence on how the wish – resulting from drive tension after secondary processing – is to be treated, determining if and how wish-fulfillment can be achieved.

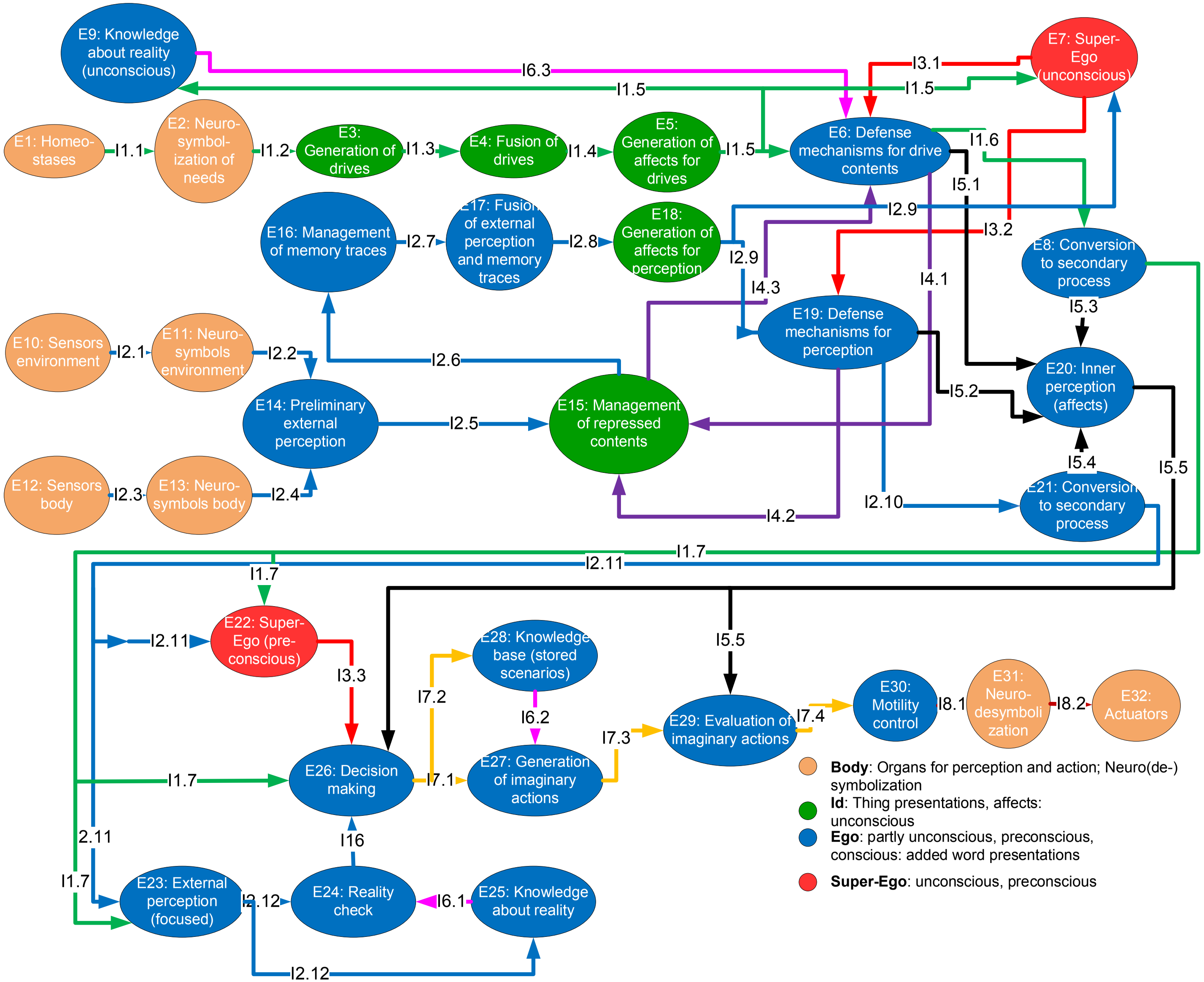

Figure 6: *Functional Model of the Psychic Apparatus*

Both, attentive outer perception (E23) and learned facts about reality (E25) affect reality checking (E24), which informs the decision unit in what way wish-fulfillment may be realistically achieved.

After drawing a pre-conscious/conscious decision that wish-fulfillment is to be aimed for (E26), with the help of memorized scenarios (E28) potential action plans are constructed (E27) and evaluated (E29). The decision for conducting a particular action plan is also influenced by feelings and affects (I5.5) resulting from inner perception (E20).

Finally, the accepted action plan is decomposed into instructions for motility control (E30), which are then neuro-desymbolized (E31), i.e. translated into physical signals to control the actuators (E32). The sensors that perceive the actual effect of the resulting actions on the body and the on the environment close the feedback loop.

5. Simulation

The results from the modeling process up to now need to be thoroughly examined and validated in practical investigations. Therefore, a simulation platform for software agents has been created at the Institute of Computer Technology, in which agents with a technical implementation of the psychoanalytical model [22], [23] are exposed to a virtual environment in which they have to survive (take decisions and conduct actions) according to the demands of their body and environment. The simulator comprises the agents as well as food sources, other creatures (fauna), and inert objects that are used for orientation. All objects of the simulated world possess basic physical properties that are computed in a physics engine (e.g. collision detection or friction). The agents are able to perceive their environment via sensors and can affect the surroundings via actuators. It receives additional information about the own body via bodily sensors. The two types of sensors together deliver information about the so-called outer world. The name comes from the fact that the information is not within the psyche, but outside – in the physical world. The information is forwarded to the current implementation of the psychoanalytical model where it causes perception particles, or drives and affects, respectively. They are represented in the form of thing presentations. Based on these conflicting forces the decision unit has to find a compromise between the demands of the body and the environment. For this reason the agent is able to think in advance – i.e. it does not actually commit an action but reasons about its potential effect and consequences on the agent itself and the environment. In a further step the thinking in advance is to be assisted by a physics engine that can give more precise statements about potential outcomes of actions.

6. Future Strategies and Conclusions

The top-down approach in psychoanalysis leads to a very complex system, however it has to be considered that it does not cover all aspects about mental functions, nor all functions of the human brain. Examples are the hormonal system or the various learning strategies[9] on different layers. The presented approach needs to be seen as a first approximation. In a next step the implemented simulation has to be investigated metrologically. This means that psychological empirical principles have to be adopted and applied to the simulation environment. In the past the development can be seen as straight-forward process with open feedback – a purely theoretical work for approaching the principles of the mental apparatus. As has already been shown with the presented European research project SENSE [4], what is left is to further develop control systems and systems that actually take decisions based an application and environment-specific approaches.

Literature

[1] Döbrich, U.; Noury, P., "ESPRIT Project NOAH – Introduction", Proceedings of the Fieldbus Conference FeT'99, 1999, S. 414-422

[2] Sauter, T.; Dietrich, D.; Kastner, W. (ed.): EIB Installation Bus System; Publicis, 2001

[3] Simo J., Benet G., Andreu G: "Embedded Video Processing for Distributed Intelligent Sensor Networks", Proceedings of the 3rd IC-SCCE, 2008.

[4] D. Bruckner, J. Kasbi, R. Velik, W. Herzner: "High-level Hierarchical Semantic Processing Framework for Smart Sensor Networks", IEEE HSI, 2008.

[5] Dietrich, D., "Evolution potentials for fieldbus systems", IEEE International Workshop on Factory Communication Systems WFCS 2000; Instituto Superior de Engenharia do Porto, Portugal, 2000

[6] E. Brainin, D. Dietrich, W. Kastner, P. Palensky, C. Roesener, "Neuro-bionic Architecture of Automation Systems – Obstacles and Challenges" 2004 IEEE Africon, 7th Africon Conference in Africa, Technology Innovation, Vol. 2, pp. 1219-1222, 2004.

[7] Luria, A., R.Pribram, K.H., "The Working Brain - An Introduction in Neuropsychology", Basic Books, 1973

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

[9] D. Dietrich, G. Fodor, G. Zucker, and D. Bruckner (Editors): "Simulating the Mind, A Technical Neuro-psychoanalytical Approach", Springer Wien New York, ISBN 978-3-211-09450-1, 2008.

[10] G. Russ, Situation Dependent Behavior in Building Automation, PhD Thesis, TU Vienna, 2003.

[9] The various learning principles developed in AI have to be adopted to the model first.

[11] Tamarit C.: Automation System Perception – First Step towards Perceptive Awareness, Dissertation Thesis, Vienna University of Technology, 2003
[12] B. Palensky (née Lorenz): From Neuro-Psychoanalysis to Cognitive and Affective Automation Systems, PhD Thesis, TU Vienna, 2008.
[13] C. Rösener: Adaptive Behavior Arbitration for Mobile Service Robots in Building Automation, PhD Thesis, TU Vienna, 2007.
[14] P. Palensky, D. Bruckner, A. Tmej, T. Deutsch: Paradox in AI - AI 2.0: The way to machine consciousness. In: CD Proceedings of IT Revolutions, 2008.
[15] M. Ulieru and R. Doursat: Emergent Engineering: A Radical Paradigm Shift, Int. J. Autonomous and Adaptive Communications Systems, 2009.
[16] A. Kirlyuk and M. Ulieru: IT Complexity Revolution: Intelligent Tools for the Globalised World Development, In: CD Proceedings of IT Revolutions, 2008.
[17] Bartuska, H, Buchsbaumer, M. F., Mehta, G. et al., Psychotherapeutische Diagnostik, Leitlinien für den neuesten Standard, Springer Verlag, 2005
[18] D. Dietrich, G. Zucker: New approach for controlling complex processes. An introduction to the 5th generation of AI, Proc. of the HSI, 2008.
[19] R. Velik: A Bionic Model for Human-like Machine Perception, PhD Thesis, TU Vienna, 2008.
[20] M. Solms, O. Turnbull, The Brain and the Inner World, Karnac/Other Press, London, England, 2002.
[21] A. Damasio, Looking for Spinoza: Joy, Sorrow, and the Feeling Brain, Harvest Books, 2003.
[22] T. Deutsch, A. Gruber, R. Lang, R. Velik: Episodic Memory for Autonomous Agents. In: Proc. of HSI, 2008.
[23] C. Rösener, R. Lang, T. Deutsch, A. Gruber, B. Palensky: Action planning model for autonomous mobile robots. In: Proceedings of INDIN, 2007.

Web-Links

[W1] http://www.youtube.com/watch?v=bCK64zsZNNs
[W2] WAP: Wiener Arbeitskreis für Psychoanalyse, http://www.psychoanalyse.org
[W3] WPV: Wiener Psychoanalytischen Vereinigung, http://www.wpv.at
[W4] http://www.neuropsa.org.uk/npsa
[W5] http://www.sfu.ac.at