8th IEEE International Conference on Industrial Informatics

INDIN 2010

July 13-16, 2010

Osaka University Nakanoshima Center, Osaka, JAPAN

IEEE

EXPO'70

Osaka Convention & Tourism Bureau

Abstraction Levels for Developing a Model of an Intelligent Decision Unit

Dietmar Dietrich†, Dietmar Bruckner†, Brit Müller†, Gerhard Zucker‡, Friederich Kupzog†
†Institute of Computer Technology
Vienna University of Technology, Austria, Europe
{dietrich, bruckner, mueller, kupzog}@ict.tuwien.ac.at
‡ Austrian Institute of Technology
Energy Department
gerhard.zucker@ait.ac.at

Abstract – **Technical processes shall be controlled more precisely and therefore require better interconnection. In order to enable machines to make qualified decisions, the interpretation of processes shall happen on a higher level of abstraction. Artificial intelligence (AI) and Cognitive Science (CS) have delivered excellent results so far, but they are still far away from being able to design intelligent units with abilities similar to those of the human mental apparatus. We show in this paper that a new approach in bionics is required using natural science to achieve a breakthrough based on information theory. We discuss the crucial aspects that have directed engineers into a dead end. And we show that new goals can be achieved, if we consider psychoanalysis and establish tight cooperation with experts from this field. The goal is the definition of a formal approach of the lowest abstract level for the mental apparatus.**

I. Motivation

Process control in the near future requires ever-increasing precision. This goes together with the resulting necessity to grasp all dependencies and interactions in a process. By doing so the process gets more complex and harder to describe formally. As an example we can look at processes in which the final decision still has to be done by a human instead of a machine: parameterizing thousands of nodes in a building, control of mobile robots in an industrial plant, airport security, security in military scenarios, and many more cannot be done by technical intelligences yet.

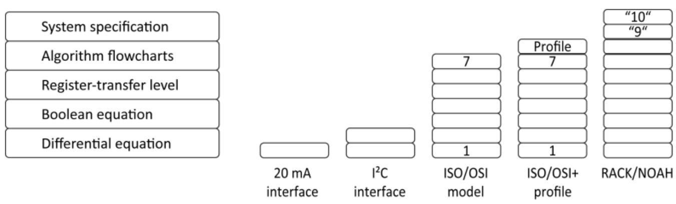

Fig. 1. Possible layer models in computer and communication technology

Another rather trivial but very real scenario shall emphasize the need for such an intelligent unit [1]. A child is in a kitchen with boiling water on the stove. Can a technical intelligence detect such a potentially dangerous scenario and take measures accordingly? If it were possible to create such a unit in mass production, reasonable pricing should become possible. Other examples are office buildings and airports: security policies require identification of persons that enter prohibited areas and identification of unattended luggage. The challenge is the separation of dangerous from safe situations.

Today, Artificial Intelligence (AI) and Cognitive Science (CS) have not solved this task satisfactorily; the routine at airports still relies on human operators that monitor surveillance cameras. We have to ask how a synergy between information theory and neuropsychoanalysis can contribute to a solution. Parts of this contribution have been given in [2], but radical new results have persuaded the authors to rework the topic, rethink the scientific situation and rephrase some of the findings.[1]

II. Layered Models and Communication Theory

In the beginning, classical electrical engineering has been the base for process automation. Electrical signals have controlled actuators and recording devices. The only relevant base for communication technology was signal theory. The complexity of processes, the diversity of the requested sequences and the economical necessity for increased automation led to a pressure to the detailed description of communication flow (protocols) and functional correlations in computers. Based on considerations that have been done in informatics and information theory engineers learned to differentiate not only between hardware and software, but also introduced several other sub layers. This abstract hierarchical differentiation of what was earlier considered only as hardware can be seen today in chip design, where the Gajski-Walker Y-Diagram describes different views on integrated circuits in the domains of behavior, structure and geometry. Software can roughly be divided into BIOS (Basic Input Output System), operating system and applications.

When looking at communication technology we find a constant increase in functional abstraction layers. In the beginning the 20-mA-interface (Fig. 1) was the only layer, later a second layer was added like, for example, in the I²C interface, then the ISO/OSI model standardized seven layers. The profiles on top of these seven layers introduce an additional eighth layer. Lately proposals for an additional increase of two more layers were made in [3, 4] to bring different fieldbusses together and to achieve virtual communication between the end users. How many communication layers will there finally be? How can a final and definite transition to the applications be achieved? Or,

[1] We would like to thank all reviewers and critics of all previous publications for their extensive statements.

when applying this thought to bionics: can principles of the mental apparatus be taken over, since humans continuously process information and also communicate it? Since this is done in far more complex ways than in technical communication systems, there will be further abstraction layers in these processes.

III. ANALYTIC CONSIDERATIONS

If methods of Artificial Intelligence (AI) don't yield the expected results we have to identify the weaknesses in these areas. We have to consider, which principles can be used and which classical approaches and methods have to be abandoned to create *intelligent systems* that meet the expected requirements.

A. Hierarchical Model

When we consider the definition of a computer as a device that manipulates, transfers and stores data, we can define the mental apparatus as described by Luria in [5] as a biological computer. If we furthermore consider the findings of psychoanalysis and assumes the functional units Id, Ego and Super-Ego as units of the upmost layer we get the layer model in Fig. 2. On the other side we have the model of an artificial computer. For now it is not relevant how many layers are considered in each model, the important thing is that both models in Fig. 2, the model of the artificial and the biological computer are functional models.

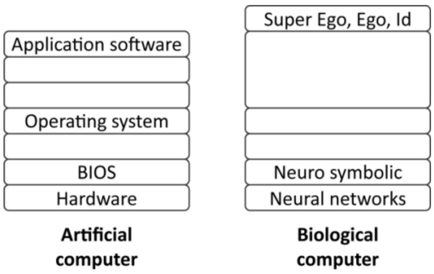

Fig. 2. Possible model of an artificial and biological computer

This enables us to use information theory, which states that layers can be exchanged by modified implementations (not by modified functionality). This means for example that the neural „hardware" can be replaced by technology of today's classical computer. In order to simulate the mental apparatus it is not necessary to use neural networks, since all functions of a neuron can be simulated by a computer. Thus it is not necessary to simulate exactly the functions of a single neuron, since it is not the exact functionality of the neuron that matters but the computation that is performed by the neural layer.

B. Top-down-Design

The achievements described in [6], in which the author separates symbolic AI, statistical AI, behavior-based AI and emotion-based AI, seen from a historical point of view, clearly show that the approaches to understand the human brain were done in bottom-up fashion. In the beginning AI first dealt with the hardware, the neurons and their interconnections and networks. We slowly worked our way up in the model and are today tackling aspects like "feelings", which, according to Fig. 2 have to be assigned to the upper layers.

An engineer who designs hardware and software for a computer discovers very soon that an optimal design can only be achieved with a clear top-down approach. Only then the optimal solution is guaranteed. Therefore the procedure for modeling has to be the following: starting from the required behavior a functional model with all inputs and outputs is derived, starting at the upmost layer in Fig. 2. All interfaces are defined unambiguously and finally the model is simulated and its behavior is verified. If the results do not meet the requirements, the model has to be modified in further iterative steps.

Such an approach for the development strongly relies on a practical consideration that has often been neglected in AI: when developing a model the experts in the field have to be consulted. An engineer is not educated to understand the function of the mental apparatus and does not have, or even know, the education of a psychoanalyst. Thus, the psycho-analysts need to define the upmost layer containing functions of the Id, Ego and Super-Ego in cooperation with the engineers. When this layer is defined, the next sub functions are defined until the lowest layer has been reached. Then it is up to the engineers to understand the model and implement it[2].

Since the layers below the layers of unconscious information processing are unknown (with the exception of the lowest, neural layer), model development is difficult. However, engineers often have to face such challenges, which shall be explained in chapter IV.

C. Behavior-based Analysis

AI and CS base their assumptions on behavior-based, psychological results, which have often been achieved by empiric research. Empiric research commonly uses statistical methods, which can derive possible dependencies. However, if we look at Fig. 2 we see a principal problem. The human being contains several physiological and mental control loops, which are more or less dependent on each other. An event occurring within the layers of such a system influences plenty of control loops. The effects that can be observed from the outside depend on the different control loops in the affected layers, between the layers and on their interconnections. Until an event can be detected from the outside it can be amplified or extinct or it can be altered to a contrary effect. The observed behavior can be the opposite of what was expected based on the event. Behavior-based statements about the mental apparatus do not help engineers at all. Copying the behavior does not allow conclusions about the functions causing the behavior. In order to create an appropriate system engineers need a functional model (and not a description of the behavior – see Fig. 3), which has been de-

[2] Of course this process has to be done iteratively between psychoanalysts and engineers – like in many developments in engineering.

signed by the according experts. When the model is simulated afterwards it can be examined using statistical (empiric) methods. If the behavior does not fit, the model has to be adapted – again by the experts for the mental apparatus, which can be again examined, and so on.

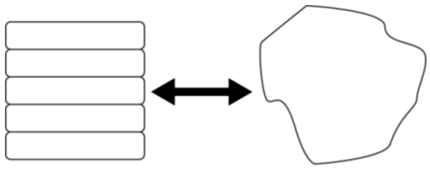

Fig. 3. Model and behavior – an abstraction

This statement is essential, since only psychoanalysis can deliver a model of the mental apparatus. Only psychoanalysis describes a *complete* functional model and its behavior without using statistical methods. Different other „schools" deliver only partial aspects, which do not suit engineers, simply because they are not educated in this way.

D. The term „Artificial Intelligence"

In Computer Engineering the term "intelligence" is widely used. In the field of digital technology a system is referred to as intelligent, if it is able to manipulate, store, and transfer data as it is defined for a computer. Psychology has another measure for intelligence. There it is investigated if and how somebody is able to fulfill particular mental operations within a given period of time. If we again compare with Fig. 2, we see that we can measure intelligence with respect to the whole system or with respect to particular layers. For the assessment of entire computer systems benchmark tests have been developed, during which particular applications (algorithms) have to be carried out as fast as possible. For the single layers of the system we could act accordingly.

Hence, the term "artificial intelligence" can hardly be measured without benchmark references. Since the goal of the model development process is to find a model of the entire human mental apparatus in order to be able to simulate it, one should be very carefully with the term of "artificial intelligence". Therefore, psychoanalysis does not define that term. Engineers as well should also avoid it.

E. The choice of the correct model

Hardware and software is basically the same thing: it is the hardware that remains or changes its binary status. For a better overview, the entity of a "computer" is abstracted and differentiated into a multi-layered, hierarchical model (Fig. 2) as described earlier (and not only in a two layered model with the layers hardware and software). This layer description also complies with Bertrand Russel's [7] monistic view of an axiomatic base with a constructive rule set. A similar abstraction process is utilized in the field of finite automata like the Mealy-device. It is composed from asynchronous RS flip flops, which are, according to the theory, separated into a logical unit and a physical unit. This allows having two abstract individual models for the two different appearances of the unit. Actually, this approach allows the development of more sophisticated circuits.

In principle, it is easy to recognize that an incorrect model leads to insufficient results. However, finding the correct model is often a stony path. Consider the following example: Previous to Kopernikus people in Europe were sure that the sun rotates around the earth because they could observe this every single day. However, people engaged in astronomy were strongly struggling. All computations of circular orbits from various planets did not become as accurate as expected. Finally, complicated models named *epicycles* were found. However, even the orbits resulting from these complex calculations were often not satisfactory and did not match reality. That contradiction between model and reality (Fig. 3) was obvious but could not be solved, neither by philosophers nor representatives of the church. Only Kopernikus did something unbelievable, he put the sun into the center of the universe. With this he disappointed the catholic clergy as well as all philosophers because their picture of the world got into trouble. Additionally, the observable behavior of the sun was different. However, with this newly adapted model the orbits of other planets simply became ellipses, which finally matched the expectations.

Observable behavior, which appears to be unambiguous, has to be seen in its context. The need for a correct, comprehensive model has already been tried to argue in section III.C and will be discussed further in the next section.

F. Interoperability of descriptions and results

Many psychological schools see their major task in helping people with problems. That is why they conduct investigations and work out descriptions of behavior and various other kinds of reflections. The goal of computer engineering, however, is different. We want to synthesize solutions. We want e. g. to take over bionic principles for which a common model is required. When reading articles like [8], where particular parts and results of various different psychological schools have been taken and put together to create a comprehensive model without checking interoperability of the sub systems (if they fit together) the computer engineer catches attention. Of course the check for interoperability cannot be conducted by engineers in this case because they lack knowledge in this field. Nevertheless it has to be conducted and the relevant experts have to be incorporated. An exaggerated argumentation: An average person living in the 15th century was not able to compare models of planetary orbits for consistency or if they possess contradictions. In the same sense engineers are overloaded when comparing various psychological schools. Additionally, they cannot be simply adopted as apparently done in [8].

G. Projection

Above, in section III.C behavior-based analysis the difference between a functional model and a description of the behavior was introduced. This abstraction is required to gain better insights into human mental processes. But, when analyzing project of machines like the CB2 [A], an additional concept has to be introduced into AI, the one of projection. Devices like the CB2 may be called "robot", but the origin of the term refers to *human-like* devices that behave like humans. If this description is reduced to solely motility or appearance, then every doll can be labeled human-like and robot. Serious comparisons with humans however need to take into account feelings, conscious and unconscious behavior. But there is yet no something like feelings in a device. Human-like robots are therefore purely wishful thinking. Everything that is said to be human-like is pure projection, nothing else. Children project something living in dolls, thus it is called human-like. Some adults project the same into a device like CB2. But it is "only" a very complex device with lots of computers and control loops. Its appearance and motility do not make it human-like.

The quality of scientific articles is often also assessed by evaluating if the researcher is in the position to consider the relevant scientific literature sources and to discuss them. Hence, if modern AI utilizes terms like emotions (emotion-based AI, see also section III.B), authors should absolutely incorporate results and findings of psychoanalysts. Who else are considered to be the experts? However, this has not yet been taken into account seriously.

H. Resolving the contradiction between the first and the second topographical model of psychoanalysis

Further review assumes another essential aspect: The first and second topographical model [9, 10] have to be treated from a different perspective. In psychoanalysis it is well known that those models cannot be aligned and protagonists have to live with contradictions. From the outside natural scientific position, when trying to create a holistic view, we need to find a way to align them in order to be able to simulate them. So, where is the problem?

In computer science data and function is strongly separated from each other. For example let us have a look at a technical model as depicted in Fig. 4. Here we have the functional computation units of a an adder and a multiplier with their data inputs and outputs.

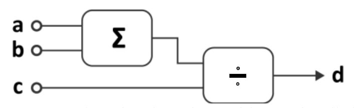

Fig. 4. Functional computational units of an adder and a divider

Of course it is not allowed to have contradicting functional specifications for the adder and the divider. However, it is allowed to have contradictions in the data! For instance, if we set the value of input c to zero, the divider divides by zero, which is mathematically undefined.

In practice, many other examples exist, e.g. two sensors measuring the same physical value at the same place but resulting in different measurement values, etc.

Technical sciences and natural sciences can cope rather well with contradictions, the only question is where they are: in the data they cannot be prevented, in the definition of functional units they must be eliminated. Hence, if we speak of contradictions within the human mental apparatus, this has to be considered *normal*, most notably if we consider that such information theoretical considerations can only be shaped with the help of computer engineering methods. In this respect the second topographical model of psychoanalysis (leaving out the content of the data) can be seen as purely functional description of the mental apparatus. This description has to be consistent. Also the first topographical model can be seen as a definition of parameters of the data, i. e. whether they are processed consciously or unconsciously. So, there is no contradiction anymore between those two topographical models. However, there is one necessary conclusion: the terms *the unconscious* and *the conscious* need to be avoided since the data can be processed, stored, and transferred *unconsciously* or *consciously* in the functional units of the *Id*, the *Ego* and the *Super-ego*. *The unconscious* and *the conscious* are not discrete functional units, they are parts of the functional units of the *Id*, the *Ego* and the *Super-ego*.

IV. NEUROPSYCHOANALYSIS

We consider a model of a system or process, where we can define the upper layer with a model "3" and the lower layer with a model "1" (as shown in Fig. 5), but not the functionality of the sub system in-between. Then engineers do the following: interfaces 1/2 and 2/3 are elaborated based on the specifications of models 1 and 3. The requirements from the two models towards model 2 build the base for the specification of model 2, for which hopefully optimal solutions can be found.

When transferring these considerations to the human mental apparatus, the depiction in Fig. 6 develops.

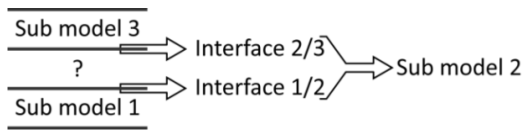

Fig. 5. System hierarchy including models and interfaces

Model 3 in Fig. 5 and 6 is represented by psychoanalysis. However, a description of psychoanalysis is required that differentiates functions from data. The requirements from psychoanalysis to model 2 need to be specified clearly. Model 1 in Fig. 5 and 6 is represented by neurology. Here, the outstanding work of [11] needs to be mentioned, in which the attempt is made to cover large portions of model 2 and to fulfill the requirements of the two adjoining models with the introduction of a neuro-symbolic model. But still mandatory questions are remaining, since [12] shows for instance that in the human brain the information channels between evolutionary grown brain structures are sometimes not fully bidirectional as expected. Another question is stability due to the necessity of learning. The human mental apparatus develops during growing up and also in adults. What does that mean for our functional model? Is this relevant for particular functions? And there are many other open questions.

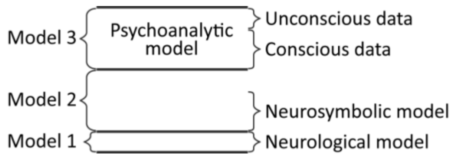

Fig. 6. Development of submodel 2

What becomes clear is: The model development process has to start on top with psychoanalysis following the details in sections III.A and III.B from "top" to "bottom". However, the specification of layer 2 (model 2 in Fig. 6) remains unknown without a full specification of layer 1.

V. THE DEVELOPED NEUROPSYCHOANALYTICAL MODEL

Based on the abovementioned considerations and the work of Luria [5], Solms [13], and Damasio [14], a model has been developed, of which parts have already been revealed in Pakistan [2] and Poland [15]. In this contribution, the model in the light of the above basic considerations is described (see Fig. 7). The lower layers are represented in the functions E1, E2, E10 - E14, E16, E17 on the input side and E31 and E32 on the output side. These functions have not been further sub divided.

It is the aim and function of the mental apparatus to synthesize the demands of three agencies. These agencies are the drive demand (E3), which bases on bodily-physiological requirements (E1), the reality demand, composed of unconscious knowledge about the facts of the outer world, its possibilities and bounds (E9), and the subjective consequences developed from perception of the outer world (E14). The third agency represents the demands of the Super-Ego (E7, E22), which is composed of socio-cultural founded bids and bans.

These three agencies and the consequences of their demands for the mental apparatus are to be comprehended. A physiological, hormonal, or other imbalance that develops in an organ (E1) triggers via neuro-symbolization of that organic requirement (E2) a drive tension (E3), whereby the drive tension is the first mental representation of bodily circumstances. Life-sustaining drives are mixed with aggressive drives (E4). Their content in the form of thing presentations is further transported (I1.4) and gets rated by an affect of more or less un-pleasure (E5). Affect together with thing presentation represents the drive, whereby the affect is the assessment thereof.

Affect and thing presentation are carried (I1.5) to the defense mechanisms (E6), which with the rules about bids and bans from the Super-Ego (E7) and under the internalized reality demands (E9) decide, if and in what form affect and thing presentation are handed over to pre-conscious and conscious processing in the Ego.

Affects and presentations which are defended through the defense mechanisms, because they are not allowed to get conscious or even pre-conscious, remain in the Id part of the mental apparatus in a container for repressed mental contents (E15).

Affects and presentations that gained access to the Ego, become processed in terms of the secondary process (E8). This means that thing presentations become additionally conjunct with word presentations. In that form thing presentations can be ordered and assessed logically, meaning they are consistent with temporal and spatial conceptions.

As has been already stated, the work of the mental apparatus is influenced by drive demands (E3) arising from bodily needs, perceptions of objects and circumstances of the outer world, i. e. the body and the environment (E14), because the outer world is represented in sensor data of two kinds, one targeted towards the environment (E10), and one solely targeted towards the own body (E12). This raw sensor data is transformed into neuro-symbols representing body and environment (E11, E13). These are the contents of unconscious perception before any kind of psychic processing or assessment. The perceived thing presentations come into contact with repressed mental content and together activate memories (E16). As a result, a combination of perception from the outer world and memory is constructed (E17). This means that not some kind of copy of the outer world is psychically processed, but a subjectively and individually assessed and associated presentation thereof. Just like the drive tensions the presentations of the outer world produce affects (E18). Presentations and affects on this level originating from outer perception and demands are similarly treated by defense mechanisms (E19) as it has been described for drive tensions. Again, perception content that passed the defense mechanisms and therefore gained access to the Ego (by whatever means), become processed in terms of the secondary process (E21).

At this level for the first time pre-conscious or conscious inner perception (E20) of drive tensions and perceived content from the outer world is possible. Inner perception includes also pre-conscious/conscious perception of feelings and affects that are related to the combined thing/word presentations.

The thing/word presentations and related affects are carried (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 pre-conscious Super-Ego in principal contains the same content as the unconscious one. Its content can be accessed consciously and has influence on the decision unit (E26). It controls how the wish – resulting from drive tension after secondary processing – is to be treated. In other words: if and how wish-fulfillment can be achieved.

Both, the attentive outer perception (E23) and the memorized facts about circumstances of reality (E25) affect the reality check (E24), which tells the decision unit what in terms of wish-fulfillment can realistically be achieved and what not – without moral assessment.

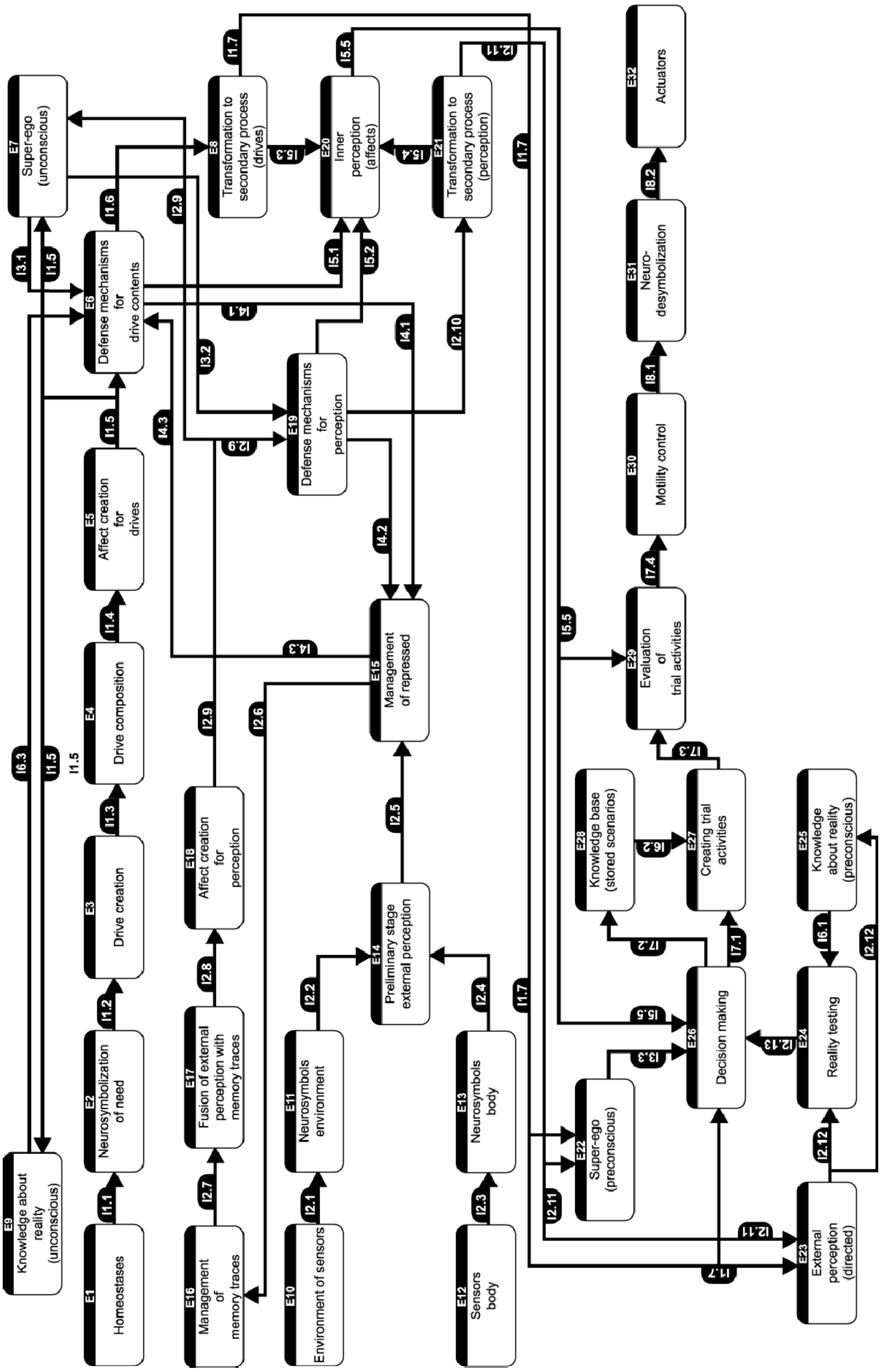

Fig. 7: Neuropsychoanalytic model [15]

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

Finally, the action plan is decomposed into instructions for motility control (E30), which are neuro-desymbolized (E31), meaning that mental content is translated into physical signals which control the actuators (E32). The feedback loop is closed via the sensors that perceive the actual affect on the own body and also the actual affect on the environment of the conducted actions.

VI. FURTHER STEPS

The model described in this paper is now in its third phase. The first, basic simulation model of a psychoanalytic agent has been presented in January 2010 [16]. The considerations above have now to be included into this simulation model, especially those of Section III.H, the combination of the first and second topographical model. A practical realization within a simulation environment is expected to be publicly presented in summer this year. It is planned that multiple simulated agents incorporate an individual, comprehensive psychoanalytic model. The memory of these agents will – in a first step – be individually pre-defined and programmed, since a learning system is not assumed here in order to avoid instabilities [11][3]. With the achieved results it will be possible to conduct simulated experiments that have to be tested on the basis of empirical psychology. Since the project has a strong basic research character, results cannot be implemented directly in practical applications, however more and more results will find application is other research initiatives like the EU-funded SENSE project [17].

REFERENCES

[1] D. Dietrich, G. Fodor, G. Zucker, and D. Bruckner (Editors): Simulating the Mind, A Technical Neuropsychoanalytical Approach, Springer, 2008.

[2] D. Dietrich, G. Zucker, D. Bruckner, B. Mueller: Bionic models for control platforms; International Conference on Frontiers of Information Technology (FIT09), Abbottabad, Pakistan, December 16~18, 2009.

[3] K. Kabitzsch, D. Dietrich, G. Pratl: LonWorks - Gewerkeübergreifende Systeme, VDE Verlag, 2002.

[4] U. Döbrich, P. Noury: ESPRIT Project NOAH – Introduction, Proceedings of the FeT, 1999, S. 414-422.

[5] A. Luria, R. Pribram: The Working Brain - An Introduction in Neuropsychology, Basic Books, 1973.

[6] B. Palensky: From Neuro-Psychoanalysis to Cognitive and Affective Automation Systems, PhD Thesis, TU Vienna, 2008.

[7] B. Russel, A. N. Whitehead: Principia Mathematica, 1913

[3] For this issue, a number of comprehensive studies still have to be undertaken.

[8] C. L. Breazeal: Designing Sociable Robots; The MIT Press, Cambridge, Massachusetts, London, England, 0-262-52431-7, 2002.

[9] Freud, S. (1915). The Unconscious. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works, 159-215

[10] Freud, S. (1923). The Ego and the Id. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIX (1923-1925): The Ego and the Id and Other Works, 1-66

[11] R. Velik: A Bionic Model for Human-like Machine Perception, PhD Thesis, TU Vienna, 2008.

[12] G. Wiest: Hierarchien in Gehirn, Geist und Verhalten; Springer Wien New York, 2009.

[13] M. Solms, O. Turnbull: The Brain and the Inner World, Other Press, London, England, 2002.

[14] A. Damasio: Looking for Spinoza: Joy, Sorrow, and the Feeling Brain, Harvest Books, 2003.

[15] D. Dietrich, R. Lang, D. Bruckner, G. Fodor, B. Müller: Limitations, Possibilities and Implications of Brain-Computer Interfaces, 3rd International IEEE Conference on Human System Interaction HSI 2010, Rzeszow, Poland, May 13-15, 2010

[16] R. Lang: “A Decision Unit for Autonomous Agents Based on the Theory of Psychoanalysis”, PhD Thesis, TU Vienna, 2010

[17] D. Bruckner, R. Velik, and G. Zucker: “Network of Cooperating Smart Sensors for Global-View Generation in Surveillance Applications”, Proceedings of the 6th IEEE INDIN, 2008.

INTERNET LINKS

[A] http://www.youtube.com/watch?v=bCK64zsZNNs