Journal of Machine Engineering, Vol. 9, No. 1, 2009

cognitive tolerance matching, self-optimization,
SOAR, cognition, intelligent production systems

Robert SCHMITT[1,2]
Carsten WAGELS[1]
Mario ISERMANN[2]

PRODUCTION OPTIMIZATION BY COGNITIVE TECHNOLOGIES

Today, value chains are considered fractionally and on the basis of simplified model assumptions. Interactions between processes, materials, means of production and individuals acting in this environment as well as the effect of changes on the product usually are not known exhaustively. In order to take corrective actions towards these deficits, self-optimizing production system technologies can be used. They provide systems that emulate the "human" ability of reaching a decision with technical architectures. The goal of these approaches is to steadily analyze and evaluate the actual status in technological as well as in organisational areas and conduct a system adaptation to alternating objectives. Central questioning in this field of research is how to survey production data in order to detect correlations of production parameters and their influence on product parameters, how to derive decisions from this knowledge and how to learn from the consequences. Application technologies capable of taking on these tasks of self-optimization to emulate intelligent behaviour are analysed. The aim is to identify the competencies of these technologies, in order to build a cognitive system architecture based on applications especially suited for each task that has to be fulfilled to emulate cognitive human decision making processes.

1. NEED AND POTENTIAL

During the last years, general conditions for the producing industry in high-wage countries have changed significantly. Due to an increased world wide competition, high quality products are requested as well as cost-efficient production techniques. Many companies are not up to these challenges in Germany and relocate their production in low cost counties. There low wages and taxes as well as scale effects of cheap mass production pretend a higher productivity. [1]

Another global trend shows that customers ask more and more for individualized products, which have to be available within a very short time. One example is the automotive industry, where customers are allowed to change the configuration of their car until a short time before delivery. As a result, successful companies have to be able to react

[1] Werkzeugmaschinenlabor WZL, RWTH Aachen University, Steinbachstr. 19, 52074 Aachen. Germany
[2] Fraunhofer Institute for Production Technology IPT, Steinbachstr. 17, 52074 Aachen, Germany

to individual customer demands quickly. Thus short throughput times, high flexibility in production planning and stable processes are mandatory. [9]

Especially production processes of high end products own complex dependencies of a high number of production parameters and their influence on the final product, so that these relations are not understood fully. This tends to result in decreasing process stability. Production processes have to be optimized constantly to achieve the desired product characteristics and tolerances. However, these constraints do not allow near-term product variations as well as the processes are rarely designed optimally, because they are often only optimized in sections. This leads to a discrete optimization of single elements in a system, without estimating their interaction concerning the characteristics of the final product.

More effective than the optimization of single process steps is the function oriented optimization of the product. This leads to the question how single objectives within a production process, like the dimensions to come up to, can be adapted dynamically. Thereby the superior objective is the function of the product, which can be achieved more efficient. If this process variation is done autonomously by the productions system, this is called self-optimization.

On the one hand, such a dynamisation of crucial process parameters will reduce costs, because single tolerances can be expanded without missing the required product characteristics. On the other hand, the flexibility of the production process in reference to changes of the product will increase significantly.

2. THE SELF-OPTIMIZING FACTORY

A self-optimizing factory is able to conduct independently changes and fundamentally optimize the production processes concerning quality, costs and throughput time. These independent variations of internal processes are called self optimization. It is defined as a repeating execution of the following actions:

- Continuous analysis of the as-is situation
- Identification of the objectives
- Adaptation of the system behaviour in order to reach the objectives [12]

Contradictory to a classical controlling circuit based on internal decisions, a self-optimizing system is able to redefine the various sub-objectives steadily and adapt the controlling process dynamically (Figure 1). Also in contradiction to a classical controlling circuit, not a single process control loop is focused, but cascading controlling circuits focusing all domains of the company that are concerned. This means, that the whole process chain is integrated into the controlling circuit on a higher level, and that information is also made available to the controlling levels above. Thus the levels of process, process chain and production control loops are also integrated into the controlling circuits. On the top level there is the controlling circuit for entrepreneurial quality. [10] On this level the superior objectives are defined and the controlling mechanisms evaluate if they are obtained. This is done by generating sub-objectives and delegating them to the controlling circuits of these levels concerned.

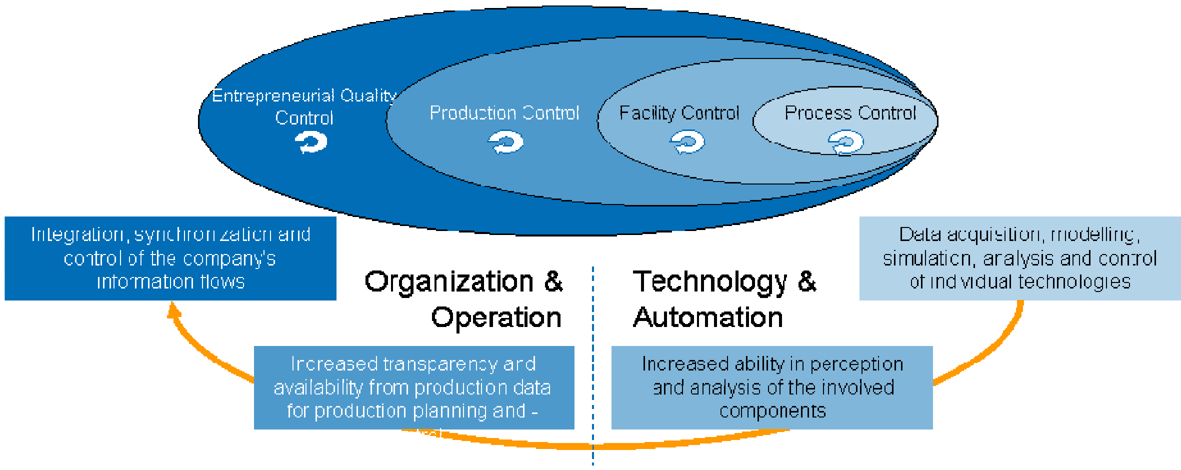

Fig. 1. Cascading controlling circuits.

As a result of the similarity to cognitive processes, self-optimization is intimately connected with the cognition science. Its research areas are cognitive systems. They consist of knowledge, a central processing unit and the surrounding situation. Cognitive systems are able to gather information from their environment, to process it in their controlling unit and to convert it into actions which have influence on the environment. [14]

Such a cognitive system is able to form the core of a self-optimizing factory. [15,16] Implementation of cognitive mechanisms on computer systems allows the exact processing of huge amounts of data and so even the analysis of complex production processes with complex multi-layer dependencies.

In the manufacturing, a cognitive, self-optimizing system is able to modify the production process by selectively changing dimensions and tolerances, in order to react to deviations of upstream process steps. The process can be stabilized and the quality of the final product can be increased. In assembly, a cognitive system can act in a similar way. By paring of individual parts and modifying actuating variables, dimension variations resulting from manufacturing can be compensated. [8]

During the inspection, all accumulated measurement data is collected. On the one hand, it can be directly sent to upstream process steps; on the other hand, the cognitive system is able to detect correlations between production and product parameters of individual parts.

Furthermore the cross process orientation of the cognitive system allows steering and optimizing the material and information flow. The individual processes are coordinated, thus stocks between process steps can be reduced and so the whole efficiency will be enhanced.

The central elements for the integration of cognitive mechanisms in a production process are a steady information flow through the whole process and an enduring optimization in the context of the continuous improvement process. [11]

The integration of a cognitive system into the whole process chain is described in figure 2. The cognitive system is a central element within the process chain. From every process step, information is gathered and processed. Solutions are elaborated and modifications are executed. Thereby the functionality of a part is the main objective of the modifications and optimizations.

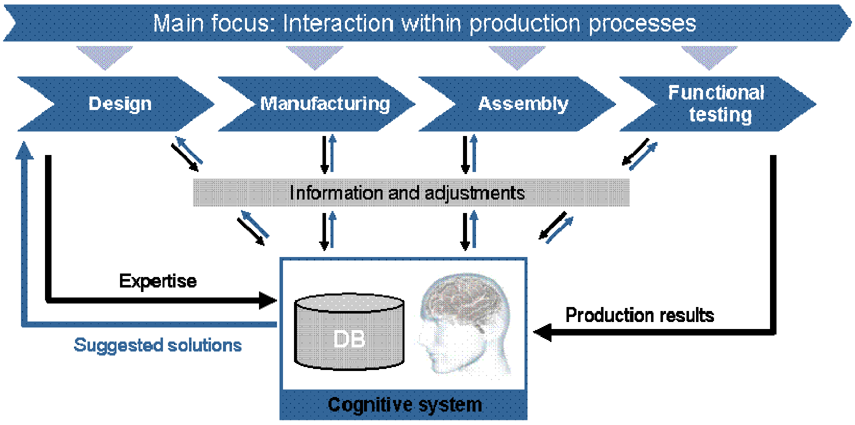

Fig. 2. The self-optimizing factory.

3. COGNITIVE SYSTEMS

Based on the academic definition, cognition can be understood as a collective term for all processes and structures that are connected to perception and identification, like thinking, memory, imagination, learning, etc.

Since the 1970s the cognitive science has developed to an interdisciplinary field of research. Its main aspects are the cognitive processes and structures of humans as well as other organisms and even technical systems (artificial intelligence). Thereby computer models and psychology are connected [6].

3.1. STEP MODEL OF INFORMATION PROCESSING

In all processes that are executed by human beings in their environment, stimuli are gathered, processed and afterwards an action is executed. This scheme can be divided into three steps of information processing.

First there is input information, which is described in perception and interpretation. Afterwards, there is the phase of information processing. The discovered information is processed and a decision to act is derived. Finally information is emitted. The information output is described by several actions that describe a concrete reaction to a stimulus [5].

Information input (Perception and interpretation): The first step of the human information processing is gathering input information. A physical stimulus is discovered or a necessary piece of information is perceived through the various sense organs. This stimulus needs a certain threshold to be perceived and categorized as important. The interpretation follows the detection of a stimulus. The stimulus is identified and compared to existing cognitive schemes [5].

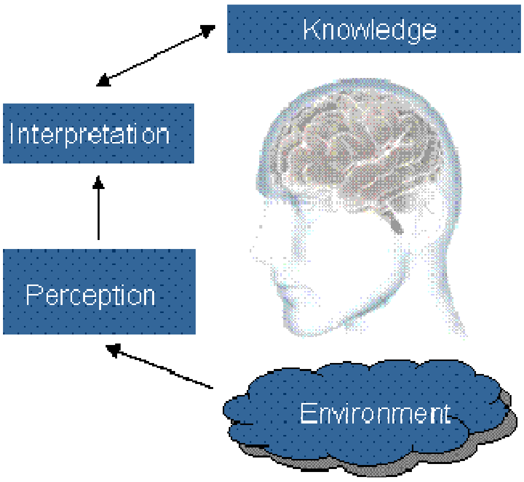

Fig. 3. Cognitive information processing

Information input (Perception and interpretation): The first step of the human information processing is gathering input information. A physical stimulus is discovered or a necessary piece of information is perceived through the various sense organs. This stimulus needs a certain threshold to be perceived and categorized as important. The interpretation follows the detection of a stimulus. The stimulus is identified and compared to existing cognitive schemes [5].

Information processing (Decision): Between the information input and the human reaction, there are a lot of mental processes. The stimulus arrives at a receptor and has to be transformed into a cognitive representation and afterwards into a reaction. During the information processing, the information is forwarded in sense of a task fulfilment. The decision is a very important aspect. The decision is a conflict solving process, which has to choose between several possible options.

Information output (Action): The information output process depends on the problem, the outer conditions and the constitution of the human being. In order to exert influence on the environment, there has to be an information output. In the context of the human cognition, information output mostly takes places with the movement of body parts or the acoustical output of speech. Both are controlled from the motion centres in the central nervous system, which are located in several brain regions and the spinal cord. [5]

3.2. COGNITIVE TECHNOLOGIES

On basis of the modelling of cognitive processes, there are several technical systems that emulate cognition with computer architectures. With SOAR, the most important one is introduced. Furthermore, artificial neural networks are discussed. An artificial neural network itself does not consist of a cognitive architecture, but has the ability of autonomous perception and learning and therefore can be utilized as part of a cognitive system acting as a self-optimizing controlling application for production systems.

Both technologies do not represent a complete cognitive system capable of carrying out every single process that is necessary for an implementation of self-optimizing controlling systems. But both of them have their own special abilities emulating the different cognitive process steps and can be combined to build a cognitive controlling system considering their special abilities. So in the following their abilities are discussed.

SOAR

SOAR (State Operator Apply Results) is a cognitive architecture based on the early systems GPS and OPS5. Intelligence is understood as the optimal achievement of objectives.

In SOAR target oriented problem solving takes place as a heuristic search in problem spaces. The search is a successive application of operators until the target situation is reached. In addition to classical planning systems, the search in the problem space is implemented in a complex decision cycle. For knowledge representation, SOAR offers two concepts in terms of a short-term memory and a long-term memory.

The long-term memory stores operative knowledge to the construction of problem spaces and control knowledge in order to steer the search processes.

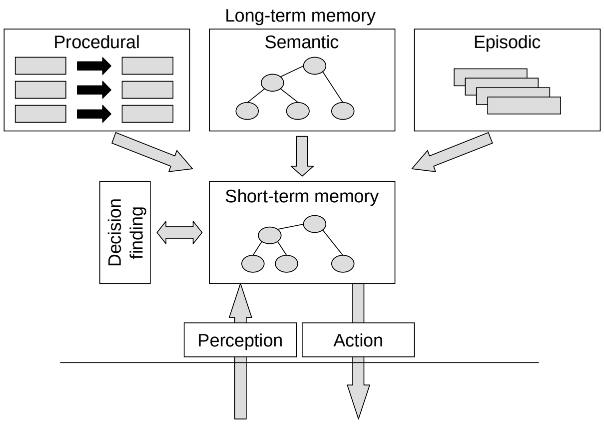

Fig. 4.The SOAR memory model

The SOAR memory model (Fig. 5) does an exact naming of the declarative memory. It is divided into the semantic and the episodic memory. Declarative knowledge is organised in so called semantic networks. The short-term memory forms the working memory. Here, all information is processed.

The uniform representation and access mechanisms and the possibility to structure the working memory in several areas result in a strong similarity to blackboards. The open arrangement of the working memory allows the adding of any modules which can use this memory or an assigned segment for information exchange and coordination.

The information processing is conducted in two phases. In the first phase, the knowledge search phase, productions of the long-term memory, which work on the working memory, fire. This process leads to the generation of new objects, which in turn activates other productions. In addition preferences, which are used in the second phase with targets, problem spaces, conditions and operators, are generated.

In the second phase, the decision procedure selects an operator by means of the actual knowledge in the short-term memory and by the aid of existing preferences and applies it to the associated problem space. By successive application of operators, either a target or a dead end will be reached. In this case, a sub target is generated in order to lead the searching process out of the dead end. If the dead end cannot be solved in this way, problem space independent mechanisms like backtracking are used. In order to avoid dead ends, a chunking learning mechanism is used. It is activated every time, when a successful way out of a dead end has been found. It generates a production rule, which consists of the entrance and the way out of the dead end. If an agent finds itself in the same situation later, the learned rule fires and the way into the dead end is avoided.

Application areas of SOAR

SOAR is already used very successfully to simulate human behaviour. Typical applications are robot controlling and the steering of artificial enemies in flight simulators for pilot training. Hereby very good results are reached, so that computer controlled aircrafts can make their own decisions and also communicate and act in teams. There are first approaches to use SOAR in special applications in the production technology [13].

Artificial Neural Networks

While SOAR acts on a symbolic level, i. e. it is based on a pre-defined concept world, artificial neural networks describe a cognitive architecture with sub symbolic information processing [4].

Artificial neurons are a technical approach of an abstract modelling to emulate the processes of a biological nerve cell. Like a biological nerve cell, artificial neurons possess input channels to detect signals in the form of input values and one output function to provide output values [2].

One integral characteristic of an artificial neural network is the ability in parallel information processing: every network entity works independently from the remaining network, thus it is not obliged to act in strict sequential actions. This would be the case in a standard architecture consisting of processor, memory and a program [14],[4].

An artificial neural network can be trained with sophisticated non-linear functions for information processing. Like in biological neural networks, the trained knowledge in artificial neural networks is represented in the weight structure of the artificial neurons. This learning differs from supervised to unsupervised learning. In case of supervised learning, the network is trained with a set of known input and corresponding output samples. With a set-actual comparison, the margin of error within the network can be identified [4].

However, in case of unsupervised learning the network itself has to generate classifications for the corresponding input signals. This is the only possible solution if there is no known output signal for the corresponding input signals [3].

Being not fully apprehensive is one of the main criticisms of artificial neural networks: Acting like a black box, one output signal is generated from the input signals without revealing anything about the inherent transfer function. This results from the complex design of neural networks accompanied with the high level of non linearity. Indeed a couple of algorithms provide the ability to generate rules regarding the behaviour of these networks, but this additional research is not fully reliable and can cause a huge amount of extra costs within a project [2].

4. COGNITIVE TOLERANCE MATCHING

Within the scope of the cluster of the research project 'Integrative Production Technology for High-Wage Countries' of RWTH Aachen University the competitiveness of the production technology of high-wage countries is to be increased enduringly. One main aspect to ensure competitiveness is self-optimizing production systems. The proposition to be proofed is that a cognitive controlled self-optimizing production system is able to act faster and more resource-efficient than a planned production system.

Self-optimizing production systems realize value stream oriented approaches with simultaneously increasing planning efficiency by transferring already acquired knowledge to similar scenarios in the production technology. In this way, new approaches for production and assembly systems are made possible. They steadily analyze and evaluate the as-is situation and reach a dynamical system adjustment regarding varying objectives.

At the moment, value stream chains are not examined over the whole process, but rather in sections. Thus a holistic modelling is impossible. So interdependencies between processes, materials, means of production and individuals acting in this environment, as well as the effect of changes on the product, are usually not known exhaustively. The correlation between variations of the production parameters and the consequences for the final product are only identified inadequately.

Within the project, comprehensive approaches in the field of coordination, planning, controlling and man-machine interaction are developed. This frame of action allows self-optimization of production systems in different directions. This shall be reached by creation and implementation of cognitive mechanisms which act on the planning and the organisational level over the whole process chain [7].

Use Case

On this approach, the RWTH Aachen University, the Fraunhofer IPT and the BMW Group defined a project to optimize the production process of the rear-axle drive in respect of its emitted acoustics. This process runs from the gear set manufacturing to the assembly and owns various very sensitive tolerances with multi-level dependencies. Within the scope of this project, Cognitive Tolerance Matching (CTM) is applied.

With Cognitive Tolerance Matching, a cognitive system for the holistic optimization of tolerance chains in a production process is developed. The objective of this system is the optimization of tolerances over the whole production process [5]. This stands for the widening of unnecessary narrow tolerances in order to save costs, as well as the more exact definition of critical tolerances in order to ensure the functionality of the final product.

The challenge lies in understanding the process, to conduct the right optimizations at the right place. To achieve this understanding, during the production process ‘Cognitive Tolerance Matching’ stores and afterwards appraises all process and product parameters, which can influence the desired product characteristic. The first step of data processing analyses the correlations between the recorded data and the desired product characteristics. Therefore statistical data mining algorithms as well as for example artificial neural networks can be used. The analysis of this data, which has been recorded over a huge number of manufactured products, allows implication of natural variation within the tolerances and also manual process variations.

The chosen rear-axle drive has a distinct influence to the acoustic inside the interior of the car and possesses very complex mechanisms of structure-borne sound generation. The structure-borne sound is emitted by the car body into the interior and can be disturbing for the customer. The acoustical profile can be specifically adjusted with the gear set. Because of partial contradictory objectives like low noise, high efficiency and high durability, this adjustment is very challenging. The car acoustic, and so also the rear-axle drive acoustic are very important product characteristics for customers.

Besides the objective to produce cars with a low noise factor, it is even more important to identify the acoustical relevant parameters in design and production to ensure, that all cars are produced with a constant acoustic without spread. A special challenge besides the pure identification of the acoustical relevant parameters is their effective modification. Thereby multi layer acoustical dependencies between the single parts of the car have to be taken into account.

Axle drives and changeable speed gearboxes like manual transmission, automatic transmission and double clutch gearboxes cause vibrations by periodic meshing. These vibrations can be amplified by tolerances in the manufacturing process chain and can be quite distracting for the customer. At axle drives, the hypoid gear set is the key element for structure-born sound generation.

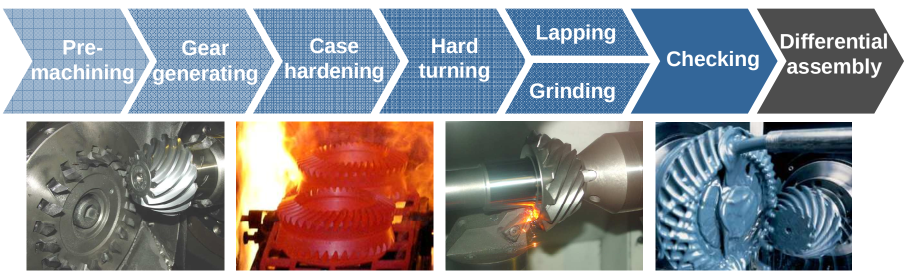

Fig. 5. Production process of a hypoid gear set

The challenge in controlling the acoustic of a rear-axle drive lies in assigning the occurring acoustical characteristics to the causing parameters (tolerances) and finally in taking the most efficient measures for error prevention.

A Cognitive Architecture

The purpose of Cognitive Tolerance Matching is to analyse the whole production process from the gear set manufacturing to the final assembly of the housing regarding the resulting acoustic and to initiate adequate optimizations. One example letting the assembly react to a derivation during the gear set manufacturing, in order to produce an acoustic optimal gear box.

With respect to the tasks mentioned one main challenge is to identify the abilities of the technologies researched to build a self-optimizing controlling system by arranging them considering their abilities of fulfilling the tasks of cognitive information processing, and to adapt the results to production systems, i. e. to the requirements of the production systems subject to their general conditions.

In the following, a combination of an artificial neural networks and SOAR is introduced as proposed solution to a cognitive system to control the production described above:

The power of SOAR is the ability to generate decisions from existing rules and validate or extend them during their application. Therewith SOAR is able to conduct a variation of manufacturing parameters, to learn from the effect of the particular application and to transform this knowledge into new rules.

So in the architecture proposed SOAR is the main element. It varies parameters representing the production, while neural networks evaluate these new generated parameter-sets subsequently. Artificial neural networks are appropriate to emulate the considered production, subject to them being trained sufficiently. A training of the neural networks is essential to emulate the action of the objects of the real production, since neural networks extract their knowledge from the application of parameters in a real system and the results obtained. That implies that a neural network is able to emulate a production, provided that there is a sufficient quantity of parameter-sets already applied representing the knowledge. However, in the application proposed neural networks are not going to make decisions, but evaluate the decision made by SOAR. That means to evaluate the vectors of production parameters and to send the results obtained back to SOAR, so that it will be possible for SOAR to learn from the results.

The detailed interaction of the combined systems is organized in the following way: Starting with a given vector of parameters and a basis set of rules SOAR conducts a variation of these parameters and sends it to the neural network. The neural network, that is assumed to be pre-trained, evaluates the parameter-sets and sends back the results obtained, i. e. the product characteristics assumed to be produced in the real production system. The results are calculated under usage of the network's knowledge about the production processes.

This procedure of varying parameters and evaluating the parameter-sets will be repeated until the results obtained by the neural network allow the conclusion, that all demands towards the product would be fulfilled in the real production. Then the parameter-set used by the neural network will also be used to produce real products. If these products made in reality fulfil all demands requested, SOAR receives a success message. The neural network's forecasting of the results to be obtained under usage of the parameters given are sufficient and so the estimation can be made, that the network was able to simulate multi-level dependencies. In the negative case, if calculated results and real results do not match, this will also be fed back to have the possibility to learn from miscalculation and to correct the rules used. If derivations from objectives occur, the following sequence of actions is suggested: After derivations are found, the network gets trained again to update the knowledge used to emulate the dependencies of the production processes. This enables SOAR to adapt to the new situation and to use its new knowledge for future decisions.

Thereby decisions are made by a systematic decision making process performed by SOAR. In doing so, improvements can be achieved systematically and faster than by using algorithms that do not consider results already achieved.

Although a specific controlling application for a real production is developed, application of the strategy introduced to other productions will also be possible, both within one single process chain and across several process chains on different organisational levels.

In the approach presented a type of production like the one introduced is focused. Experiments showed that neural nets are able to detect and to deploy the dependencies of single processes as well as simplified process chains. Experiments also showed that SOAR is able to make decision based on existing rules. Challenges of recent research are to combine those technologies subject to build a real-time capable controlling system acting like described in this paper.

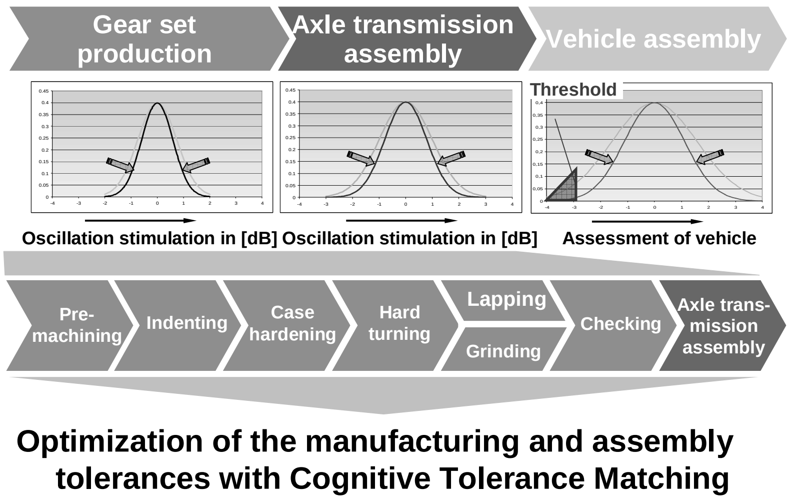

Fig. 6. Implementation of CTM in a production process.

Figure 6 shows the idea of ‘Cognitive Tolerance Matching’ in the whole production process. Information from manufacturing, assembly and usage is sent back to the design, to change the important product parameters. Furthermore, Cognitive Tolerance Matching connects the dedicated production steps and enables them to react to process derivations immediately. So, CTM contributes to the continuous improvement process in the production and supports reaching the superior product objectives, even with derivations in one or more production steps. Not the tolerance orientated production, but the functional orientated production forms the main goal of Cognitive Tolerance Matching. In this way, the perceived quality can be stabilized and even increased.

5. SUMMARY

The producing industry in high wage countries is more and more under pressure by virtue of global competitors. To keep the economically important production in these countries, an exact understanding of the processes is mandatory.

The self-optimizing factory supports the production technology in facing these requirements. It allows the fast and stable adaptation to conditions changed, supports the man-machine cooperation and allows the cost efficient production in a turbulent environment. Therefore it uses production systems, whose components dispose of their own perception, knowledge, planning and learning ability.

The approach presented in this paper deals with a cognitive controlling system for production systems, whose intelligent combination of SOAR and neural networks enables it to adapt to changing conditions fast and efficient.

Besides this approach, there are actually other promising research and development projects regarding the effective and efficient application of cognitive systems in production technology. By now, the cognitive science owns a long tradition and cognitive architectures achieved great results in other fields of applications, for example robot controlling. In the production technology, theses systems are standing shortly before their industrial application.

ACKNOWLEDGMENT

The depicted research has been funded by the German Research Foundation DFG as part of the Cluster of Excellence ‘Integrative Production Technology for High-Wage Countries’.

REFERENCES

[1] PFEIFER T. ,SCHMITT R., *Autonome Produktionszellen. Komplexe Produktionsprozesse flexibel automatisieren,* Springer Verlag, Berlin/Heidelberg, 2006.
[2] MIELKE A.: *Neuronale Netze*; 1999.
[3] HOFFMANN N., *Kleines Handbuch Neuronale Netze: Anwendungsorientiertes Wissen zum Lernen und Nachschlagen*; Vieweg, Braunschweig/Wiesbaden, 1993.
[4] KRATZER K., *Neuronale Netze*, Carl Hanser, München/Wien, 1990.
[5] LUCZAK H., *Arbeitswissenschaften*, 2. Aufl., Springer, Berlin, 1998.
[6] PUTZER H., *Ein uniformer Architekturansatz für kognitive Systeme und seine Umsetzung in ein operatives Framework,* Dr. Köster Verlag, Berlin, 2004.
[7] SCHMITT R., ISERMANN M., WAGELS C., *Cognitive Tolerance Matching*, Proceedings of the 6th CIRP International Seminar on Computation in Manufacturing Engineering (ICME), Naples, Italy, 2008.

[8] SCHMITT R., DAMM B., PAVIM A., *SeIf-Optimised Assembly of a DPSS Laser through Sensor Fusion*, 6th CIRP International Seminar, Naples, Italy, 2008.
[9] SCHMITT R., HAMMERS C., *Governing the Process Chain of Product Development with an enhanced Quality Gate Approach*, CIRP Journal of Manufacturing Science and Technology, 2009, Vol. 1, No. 3, 206-211.
[10] SCHMITT R, KRISTES D, BEAUJEAN P, *Entrepreneurial Quality Management – A new Definition of Quality*, IEEE International Engineering Management Conference, New Jersey, 2008, 275-280.
[11] SCHMITT R, SCHARRENBERG C, *Planning, control and improvement of cross-site production process chains*, Proc. 40th CIRP International Manufacturing Systems Seminar, Liverpool, Great Britain, 2007.
[12] SCHMITT R., BEAUJEAN P., *Selbstoptimierende Produktionssysteme*, ZWF - Zeitschrift für wirtschaftlichen Fabrikbetrieb, 2007, Vol. 102, No. 9, 520-524.
[13] Soar Technology: www.soartech.com, 01.12.2008.
[14] STROHNER H., *Kognitive Systeme. Eine Einführung in die Kognitionswissenschaft*, Westdeutscher Verlag, Opladen, 1995.
[15] ZÄH M. F. et al., *The Cognitive Factory*, ElMARAGHY, H. A. (Hrsg.): Changeable and Reconfigurable Manufacturing Systems, Springer, Berlin/Heidelberg, 2008.
[16] ZÄH M. F. et al., *Kognitive Produktionssysteme – auf dem Weg zur intelligenten Fabrik der Zukunft*, ZWF-Zeitschrift für wirtschaftlichen Fabrikbetrieb, 2007, Vol. 102, No. 9, 525-530.