We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



6

Motivation in Embodied Intelligence

Janusz A. Starzyk
Ohio University
U.S.A.

1. Introduction

Before artificial intelligence set its mind on developing abstract intelligent agents which can think, Alan Turing suggested training embodied machines equipped with sensors and actuators to accomplish intelligent tasks like understanding spoken English (Turing, 1950). Looking at intelligence from a different perspective, philosopher, and neuroscientist Francisco Varela (Maturana & Varela, 1980), (Varela et al., 1992) proposed the embodied philosophy of living systems which argues that human cognition can only be understood in terms of the human body and the physical environment with which it interacts. What may seem to be a revelation from a historical perspective, early robots built on cybernetic principles demonstrated goal-seeking behavior, homeostasis (the ability to keep parameters within prescribed ranges), and learning abilities (Walter, 1951), (Walter, 1953). These were precursors for embodied intelligence. Perhaps the most influential figure in developing embodied intelligence as a methodology to design intelligent machines is Rodney Brooks. He suggested the design of intelligent machines through interaction with the environment driven by perception and action, rather than by a prespecified algorithm (Brooks, 1991a).

Like Hans Moravec before him (Moravec, 1984), Brooks suggested that locomotion and vision are fundamental for natural intelligence. He also observed that the environment is its best model and that representation is the wrong "unit of abstraction". These simple observations revolutionized the way people think about intelligent machines and created a field of research called "embodied intelligence". The growth of interest in embodied intelligence that followed Brooks' works can be compared to the increase in research activities in artificial intelligence that followed the famous Dartmouth Conference of 1956 (McCarthy et al., 1955) or the revival of neural network research in the 1980s. His approach revived the field of autonomous robots, but as robotics thrived, research on embodied intelligence started to concentrate on the commercial aspects of robots with a lot of effort spent on embodiment and a little on intelligence.

The open question remains: how to continue on the path to machine intelligence? Today, once again, artificial intelligence research is focused on specialized problems, such as ways to represent knowledge, natural language and scene understanding, semantic cognition, question answering, associative memories or various forms or reinforcement learning. In recent years, the term "general artificial intelligence" was coined as something new, incorrectly implying that the original idea of AI was something less than to develop a natural intelligence.

Brooks decided to build intelligent autonomous creatures that work in a dynamically changing environment. He pointed out that he is not interested in finding how humans work, nor in philosophical implications of creatures he creates. He let them find their own niche to operate in. Although he would like humans to perceive these creatures as intelligent, he does not define what this would mean. He would like these creatures to be able to adapt to changes in the environment by gradual changes in their capabilities. Each creature should have a purpose of being; it should maintain and pursue multiple goals, choosing which goal to implement based on the environmental conditions. In addition, the complexity of a creature's behavior would reflect the complexity of the environment in which it operates rather than its own.

Proposed by Brooks, subsumption architecture leads to independent sensory-motor control structures that work concurrently and are designed such that lower level skills are subsumed by the higher levels. Thus multiple parallel sensory-motor paths must be implemented to control the creature's behaviour. He argues that no central control or representation is needed. Instead individual robot skills are built layer after layer each one composed of a simple data driven finite state machine with no central control.

Brooks seems to reject the connectionist (and implicitly neural network) approach. The finite state machines he uses to control his creatures must be explicitly programmed to perform certain actions. However, this explicit engineering approach that works successfully on very low levels of subsumption architecture does not have a natural mechanism for self-organization from which higher level skills could evolve. Machine learning, which may be a critical element of intelligence, is almost left out of consideration. Indeed, the only learning that takes place in embodied agents is based on simple neural network structures. But years of development of classical neural networks failed to deliver acceptable forms of learning due to the catastrophic interference observed in generic neural networks (McCloskey & Cohen, 1989). Yet, in my opinion, learning distinguishes the intelligent from the unintelligent. Thus, subsumption architecture may be a clever way to design autonomous robots with reactive control, but it is not a mechanism that may scale up to human level intelligence. I claim that many years after Moravec's article, subsumption architecture has still failed to solve fundamental problems of embodied intelligence and needs a major revamp.

Brooks requires that machine uses multiple, data driven, parallel processing mechanisms that control machine's behavior. Yet, he clearly differentiates his approach from this of neural networks. He claims that there is no obvious way to assign the credit or blame in neural networks for a correct or incorrect action. He pointed out that the most successful learning techniques for embodied robots use reinforcement learning algorithms (like Q-learning) rather than parallel processing neural networks. He stressed dense connectivity of neural networks that are in striking contrast to his system of loosely connected processes.

By rejecting the connectionist approach and self-organization of machine architecture, Brooks denied his subsumption architecture the flexibility to integrate evolved lower level functions into more complex levels without explicit interference of a human designer. From a system engineering point of view, each subsequent step in system complexity requires exponentially harder design effort and understanding of what the creature can do and how it does it. Yet as Brooks observed, this was not the case in nature. It took nature over 3 billion years to create insects from the primordial soup, but it took only 200 million more years to create mammals, and only 15 million years for the transformation of great apes to

modern man about 3 million years ago, with all major developments of the civilized world within the last 10,000 years. It seems that in nature it is easier to append a primitive brain to create a complex brain capable of abstract thought, than it is to learn locomotion and survival skills in primitive brains. While this may justify an approach in which a machine's reflexes are developed first, the lack of a mechanism to add complexity at a low design cost is a major problem that cannot be left to chance.
Brooks rightfully indicated that development of intelligence should proceed in a bottom-up fashion from simpler to more complex skills, and that the skills should be tested in the real environment. He rightfully criticized the symbolic manipulation approach for requiring that a complete world model is built before it can be used. He also rejected knowledge representation as ungrounded. However, instead of proposing an approach that bridges the gap between processing raw sensory and motor signals, symbolic knowledge representation and higher level manipulation of symbols, he assumed a constructionist approach with no hint of how to develop natural learning. This denying the need for representation was criticized by Steels (Steels, 2003), who pointed out that representations are internal conceptualizations of the real world and thus ought to be acceptable to the embodied intelligence idea. So, in spite of its great success in building creatures that can move in a changing environment, subsumption architecture failed to create foundations for intelligence. To paraphrase Brooks' own words - the last seventeen years have seen the discipline coasting on inertia more than anything else.

In this chapter, I will present a path for further development of the embodied intelligence idea. First, I will directly address the issue of intelligence. The problem with Brooks' approach is not that he did not define intelligence, leaving it to philosophers, but that he accepted any autonomous behaviour in a natural environment as intelligent. While it is true that survival-related tasks form a necessary basis for development of intelligence, they alone do not constitute one. Is an amoeba intelligent? How about virus or bacteria? If we expect an intelligent behaviour, we need to define one. Instead of defining embodied intelligence, Brooks wants to design creatures that are seen as intelligent by humans. Still, he knows very well that a complex behaviour may result from a very simple control process. So how will he decide if an agent is intelligent? In fact, he is not interested in designing intelligent agents but instead in building working autonomous robots. Yet he claims that those reactive machines are intelligent.

Why might this be important? For a number of years in embodied intelligence, process efficiency dominated over generality. The principle of cheap design in building autonomous agents promoted by Pfeifer (Pfeifer & Bongard, 2007) supports this philosophy. It is cheaper and more cost effective to design a robot for a specific task than it is to design one that can do several tasks, and even more so than to design one that can actually make its own decisions. A computer can compute many times faster and more accurately than man, but it takes a human to understand the results. A machine can translate foreign speech, but it takes a human to make sense out of it. Thus there is a danger of using the principle of chip design to design a robot with no intelligence and call it intelligent as long as it does its job. This must not happen if we want to continue on the path to build more and more intelligent machines. So the question is what traits of embodied intelligence development must really be stressed, and where must the design effort concentrate?

2. Design Principles for Embodied Intelligence

The principles of designing robots based on the embodied intelligence idea were first described by Brooks (Brooks, 1991b) and were characterized through several assumptions that would facilitate development of embodied agents.

The first assumption was that the agents develop in a changing environment which they can manipulate through their actions and perceive through their senses. An important assumption was that there was no need to build a model for the environment; instead we could use the environment the way it is. These assumptions constrain the dynamics of agent-environment interaction. Based on Wehner's work (Wehner, 1987), Brooks suggested that evolutionary development led to the right form of interaction between sensory inputs and motor control provided by the nervous system. This led him to a design principle based on an ecological balance that must exist between the amount of information received, the processing capability and the complexity of the motor control.

Brooks rejects the need for explicit representations of environment or goals within the machine. Instead he uses active-constructive representations that permit manipulation of the environment based on graphically represented maps of environments. His statement that he does not represent the environment may be misleading. Just saying that this representation is different than traditional AI representation is not enough – a robot builds and maintains representations of the world. The fact that instead of planning ahead what to do next, an iterative map is used does not change the fact that some form of environment representation is needed. A local marker telling the robot where he is with respect to the map is also a form of environment representation.

Additional principles of designing embodied intelligence were characterized by Rolf Pfeifer (Pfeifer & Bongard, 2007) and include:

1) Principles of cheap design and redundancy. According to these principles design must be parsimonious and redundant. This means that by exploiting an ecological niche design can be simplified, while redundancy requires functionality overlap between different subsystems. Although these principles were not explicitly stated in Brooks' work, he stipulated them in his description of the design process.

2) Principle of parallel, loosely-coupled processes. This requires that intelligence emerges from interactions of lower level processes with the environment. This principle was in fact a foundation of internal organization of subsumption architecture based on Brooks' ideas and led to implementations of embodied agents that integrated many reactive sensory-motor coordination circuits using finite state machine architectures.

3) The value principle. This principle stands out among those adopted by Pfeifer as the one that tells a robot what is good for it. The agent may use this principle to decide what to do in a particular situation. In Brooks' work this is decided by competing goals but the goals are predetermined by a designer, and deciding which goal to pursue is also preset.

It was demonstrated that subsumption systems based on embodied intelligence ideas can anticipate changes in the environment and have expectations, can plan (Agre & Chapman, 1990), can have goals (Maes, 1990) and can do all of this without central control or symbolic representations.

In Brooks' article (Brooks, 1991b), an important issue related to learning in the subsumption architecture remains unsolved: how to develop methods for new behaviors, and change the existing behaviors. Brooks indicated that the performance of a robot might be improved by

adapting the ways in which behaviors change as a result of experience, however he does not say how this might be accomplished. He claims that thought and consciousness will emerge from interaction with the environment. While such a general statement is definitely true, based on nature's success in creating people who think and are conscious, there is no indication of how these may emerge in the subsumption architecture.

Pfeifer indicated that by allowing an agent to develop its own behaviors rather than having them programmed, additional properties may emerge (Pfeifer & Bongard, 2007). Although, unlike Brooks, Pfeifer admits that learning is an essential part of intelligence, he dismisses successes of machine learning fields as "almost entirely disembodied" and therefore not interesting. In addition he seems to deny the possibility of building embodied intelligence in the virtual world, and instead points out the necessity to bring it up entirely in mechanical robots. Yet there is nothing in the concept of embodied intelligence that precludes existence of a virtual embodied agent, as long as it has well-defined sensors and actuators. A virtual agent will be situated in a dynamically changing environment. Such an agent will perceive its environment through its sensors and act on it in a way similar to a robot that acts in the real world, and such an agent may do this in an intelligent way. In fact, considering the significant cost and design effort of building and maintaining robots, virtual agents should be the first rather than the last choice to develop ideas of embodied intelligence. And yes, development of good ideas and structural organization principles of signal processing elements in intelligent machines are what we need to solve the intelligence puzzle.

One of the motivations that Pfeifer uses in support of a developmental approach to cognition is the ontogenetic development of humans from children to adults, and he would like to see some form of implementation of the physical growth process. I see no such need, as a child may fully develop psychologically without the physical growth of its body. It's the brain of a child that needs to develop by experiencing the world, and the brain development is accomplished by learning proper behaviors rather than by a physical growth. In fact, the opposite may be true regarding topological complexity of the networks of neurons in the brain, as the brain of a young child has many more neural connections and therefore may have a higher ability to learn than the brain of an adult.

Pfeifer is right when he suggests that representing lower level attractor states as symbols provides a grounded way of bottom-up building of cognitive systems. This is in contrast to earlier views by Brooks, who denied that symbol manipulation may play a useful role in development of embodied intelligence. The symbols used in this bottom-up representation building are known only to the machine that holds them and cannot be explicitly defined and entered from outside (for instance by a programmer). Thus they are grounded in the machine's way to perceive and history of interactions with the environment.

Pfeifer acknowledges that the value system in embodied intelligence is murky to a similar degree as it is in biology, psychology or artificial intelligence. However, he states that the value is in the head of the designer rather than in the head of an agent. This approach to value learning is acceptable only for simple reactive systems that require external reinforcement to learn values and may not be sufficient for intelligent systems.

In reinforcement learning (Sutton, 1984), values are either associated with the machine's states or with activation of neurons in neural network implementation. However, state-based value learning is useful only for the simplest systems with a small number of states. The learning effort does not scale well with the number of states. If a system uses neurons

to learn and control its operation, then its number of states grows exponentially with the number of neurons and learning the values associated with all these states is difficult. In addition, a system that uses only external reinforcement to learn its values suffers from the credit assignment problem where credit or blame must be assigned to various parts of the system for an action that resulted in a reward or punishment (Sutton, 1984) , (Fu & Anderson, 2006).

Optimal decision making of human activities in a complex environment was rendered intractable by reinforcement learning. To remedy this deficiency of reinforcement learning, a hierarchical organization of complex activities was proposed (Currie, 1991). Expecting that a hierarchical system will improve reinforcement, Singh analyzed the case in which a manager selects its own sub-managers (Singh, 1992) who are responsible for their subtasks. Sub-managers had to learn their operation and their system of values. In a similar effort, Dayan (Dayan & Hinton, 1993) developed a system in which a hierarchy of managers was used to improve the reinforcement learning rate. It was demonstrated (Parr R. & Russell, 1998) that dividing a task into simpler tasks in reinforcement learning significantly improves learning efficiency. Based on these ideas, Dietterich used decomposition of the Markov decision process and developed a hierarchical approach to reinforcement learning (Dietterich, 2000). This divide and conquer approach requires evaluation of internal states of the machine and close supervision by a designer. In its extreme case of controlling each step, it will converge toward a supervised learning. Such a system is incapable of setting its own system of values.

A fundamental question that Pfeifer asked in his book (Pfeifer & Bongard, 2007) is what motivates an agent to do anything, and in particular, to enhance its own complexity. What drives an agent to explore the environment and learn ways to effectively interact with it? According to Pfeifer, an agent's motivation should emerge from the developmental process. He called this the "motivated complexity" principle. But isn't this like the chicken and egg problem? An agent must have a motivation to learn (and therefore to develop into a complex being), while at the same time, its motivation must emerge from this same development. Another idea for handling the motivation problem was presented by Steels (Steels, 2004), where he suggested equipping an agent with self-motivation that he calls the "autotelic principle". According to this principle the idea of "flow" experienced by some people when they perform their expert activity well would be used as motivation to accomplish even more complex tasks. However, no mechanism was proposed to identify "flow" in a machine or to implement the flow as a driving force for learning.

Many people in the embodied intelligence area ask (Steels, 2007) – where do we go now? In spite of many successes of embodied intelligence, fundamental problems of intelligence still remain unanswered. So it is quite surprising that the suggestion put forth by Pfeifer and Bongard (Pfeifer & Bongard, 2007) is to concentrate on advancements of robotic technology like building artificial skin or muscles. While this may be important for development of robots, it diverts attention from developing intelligence.

I hope that this discussion will help to bring focus back to the critical issues for understanding and developing intelligence. In the next few sections I will show how an agent may develop and maintain its system of values that controls its behavior. Such values are directly related to higher level goals and are only partially controlled by the environment. Higher level goals are established and their values learned by the machine. The machine is motivated to accomplish goals by the way it interacts with the environment.

3. Intelligence

In his seminal paper (Brooks, 1991b), Brooks pointed out that it does not matter what is intelligence and what is environmental interaction. Instead he stressed the utility of an agent's interaction with the environment and determined intelligence through the dynamics of this interaction. While this assumption helped to simplify the design of intelligent robots and justified a bottom-up approach to building intelligent machines, it also introduced a dangerous possibility of confusing a complex behavior with synonyms of intelligence. The question of intelligence is an important one if one wants to design an intelligent machine.
There is no universal agreement about how to define intelligence. However, there is a good understanding of what an intelligent agent (biological, mechanical or virtual) must be capable of. Scientists list such capabilities as abstract thinking, reasoning, planning, problem solving, intuition, creativity, consciousness, emotion, learning, comprehension, memory and motor skills as traits of intelligence. They use various tests and intelligence measures to compare levels of intelligence and differentiate between the intelligence of humans and nonhuman animals. In fact, passing various tests for (human level) intelligence was used as a substitute for its definition. Complex skills and behaviors were used to define how intelligence manifests itself. This skill based approach was inconsistent, because once a machine that was obviously not intelligent satisfied one test, another test was used in its place. This was a result of poor understanding of what is needed to create intelligence.

3.1 Definition of embodied intelligence

Existing definitions of intelligence focus on describing the properties of the mind rather than describing the mind itself. It is like defining a TV set not by how it is built and how it works but by what it does. Yet in order to design a mind we must agree on what we are designing. Perhaps driven by similar needs John Steward defined cognitive systems as follows (Stewart, 1993):

Definition: A system is cognitive if and only if sensory inputs serve to trigger actions in a specific way, so as to satisfy a viability constraint.

In a similar effort I propose an arbitrary and utilitarian definition of intelligence with the aim to present a set of principles and mechanisms from which necessary traits of intelligence can be derived. I hope that this definition is general enough to characterize agents of various levels of intelligence including human. To avoid a general discussion on intelligence I will utilize this definition to design embodied agents suggested by Brooks (Brooks, 1991b) and described in more detail by Pfeifer (Pfeifer, 1999).

Definition: Embodied intelligence (EI) is defined as a mechanism that learns how to survive in a hostile environment.

A mechanism in this definition applies to all forms of embodied intelligence, including biological, mechanical or virtual agents with fixed or variable embodiment, and fixed or variable sensors and actuators. Implied in this definition is that EI interacts with the environment and that the results of actions are perceived by its sensors. Also implied is that the environment is hostile to EI so that EI has to learn how to survive. This hostility of environment symbolizes all forms of pains that EI may suffer – whether it is an act of open hostility or simply scarcity of resources needed for the survival of EI. The important fact is that the hostility is persistent. For example, battery power level is a persistent threat for an agent requiring it. Gradually the energy level goes down, and unless the EI replenishes its energy, the perceived discomfort from its energy level sensor will increase.

Hostile stimulation that comes from the environment towards EI is necessary for it to acquire knowledge, develop environment related skills, build models of the environment and its embodiment, explore and learn successful actions, create its value system and goals, and grow in sophistication. Thus perpetual hostility of environment will be the foundation for learning, goal creation, planning, thinking, and problem solving. In advanced forms of EI it will also lead to intuition, consciousness, and emotions. Eventually all forms and levels of intelligence can be considered under the proposed definition of EI.

A critical element of the EI definition is learning. Thus an agent that knows how to survive in a hostile environment but cannot learn new skills is not intelligent. This will help to draw the line between developmental systems that learn from those that do not and perhaps will help to differentiate intelligent and non-intelligent animals. In this definition, purely reactive systems that do not learn are not intelligent, even if they exhibit complex behavior. A system must maintain its learning capability for us to continue calling it an intelligent system.

Notice that this definition of EI clearly differentiates knowledge from intelligence. While knowledge represents the acquired set of skills and information about the environment, intelligence requires the ability to acquire knowledge. Knowledge is a byproduct of learning, thus it is not necessary to include a pre-existing knowledge base in the machine memory. In turn, learning requires associative memories capable of storing spatio-temporal information acquired over various time scales. Learning to survive requires not only memory but its management, so that only the important memories are retained. Learning also requires the ability to associate the sensory and motor signals, so that action outcomes can be linked with causes.

3.2 Embodiment and intelligence

Intelligence cannot develop without an embodiment or interaction with the environment. Through embodiment, intelligent agents carry out motor actions and affect the environment. The response of the environment is registered through sensors implanted in the embodiment. At the same time the embodiment is a part of the environment that can be perceived, modelled and learned by intelligence. Properties of the motors and sensors, their status and limitations can also be studied, modelled and understood by intelligent agents.

The intelligence core interacts with the environment through its embodiment, as shown in Fig. 1. This interaction can be viewed as closed-loop sensory-motor coordination. The embodiment does not have to be constant nor physically attached to the rest of a body that contains the intelligence core (brain). The boundaries between embodiment and the environment change during the interaction which modifies the intelligent agents' self-determination. Because of the dynamically changing boundaries, the definition of embodiment contains elements of indetermination.

Definition: Embodiment of EI is a mechanism under the control of the intelligence core that contains sensors and actuators connected to the core through communication channels.

A first consequence of this definition is that the mechanism under control may change. When the embodiment changes, the way that the embodiment works and the intelligent agent interacts with the environment will be affected. Second, embodiment does not have to be permanently attached to the intelligence core in order to play its role of facilitating sensory-motor interaction with the rest of the environment. For instance, if we operate a

machine (drive a car, use a keyboard, play tennis), our embodiment dynamics can be learned and associated with our actions to an extent that reduces the distinction between the dynamics of our own body and the dynamics of our body operating in tandem with the machine. Likewise, artificially enhanced senses can be perceived and characterized as our own senses (e.g. glasses that improve our vision or a hearing aid that improves our hearing). Another example of sensory extension could be an electronic implant stimulating the brain of a blind person to provide visual information. Third, not all sensors and actuators have to probe and act on the environment external to the body. While those that do, allow the EI to interact with the external environment, internal sensors and actuators support the embodiment. When its body temperature rises, a machine may activate an internal cooling mechanism. When an animal is threatened, its heart beats faster in preparation for a fight or an escape. The body experiences internal pain that communicates a potential threat. Thus a flow of signals though embodiment is as shown in Fig. 1.

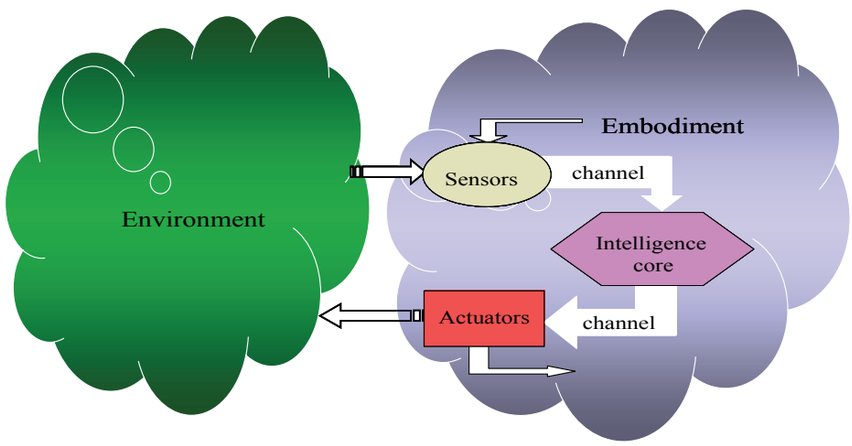


Fig. 1. Intelligence core with its embodiment and environment

Extended embodiment does not have to be of a physical (mechanical) nature. It could be in the form of remote control of tools in a distant surgery procedure or monitoring the Martian landscape through mobile cameras. It could also be our remote presence at the soccer game through received TV images or our voice message delivered through a speakerphone to a group of people at a teleconference.

An extended embodiment of intelligence also may come in the form of organizations and their internal working mechanisms and procedures. A general directing troops on a battle field feels a similar power of moving armies as a crane operator feels the mechanical power of the machine that he operates. In a similar way, a president feels the power of his address to his nation and the large impact it makes on people's lives.

This extended embodiment enhances EI's ability to interact with the environment and thus its ability to grow in complexity, skills and effectiveness. If the President learns how to address the nation, his abilities and skills to affect the environment grow differently than that of a woman in Darfur trying to save her child from violence.

Our knowledge of embodiment properties is a key to its proper use in interaction with the environment. We rely on this knowledge to plan our actions and predict the responses from

the environment. A change in the way our embodiment implements desired actions or perceives responses from the environment introduces uncertainty into our behavior and may lead to confusion and less than optimal decision making. If a car's controls were suddenly reversed during operation, a user would require some adaptation time to adjust to the new situation and might not be able to do it before crashing. Therefore, what we learn about our environment and our ability to change this environment is affected not only by our intelligence (ability to learn, understand, represent, analyze and plan) but by correct perception of our embodiment as well.

3.3 Designing an embodied intelligence

Learning is an active process. EI acquires information about its environment through sensors and interacts with it by sensory-motor coordination. The motor neurons fire in response to excitations according to desired actions associated with the perceived situation.
Learning which actions are desirable and which are not makes the learning agent more fit to survive in a hostile environment. There are several means of adapting to the environment that an agent can use to survive: evolutionary - by using the natural selection of those agents that are most fit, developing new motor skills like sweating in the hot weather or new sensors like cell sensitivity to light; and cognitive - by learning, using memory and associations, performing pattern recognition, representation building, and implementing goals. Here we address only the latter form of adaptation for the development of EI as the one we associate with an agent's intelligence. Another important form of intelligence - group intelligence - is left for future consideration, as it depends on the individual intelligence of the group members.
All spatio-temporal patterns that we experience during a lifetime underlie our knowledge, and lead to internal models of the environment. The patterns have features on various abstraction levels, and relations between these features are learned and remembered. Abstract representations are also built to represent motor actions and skills. The perceptual objects that a person can recognize, the relations among the objects, and the skills that he has are all represented in his memory. The memory is episodic and associative. It is distributed, redundant, and parallel, short term or long term. Various parts of memory are interconnected and interact in real time.
Another critical aspect of human brain development is self-organization. By self-organizing their interconnections, neurons quickly create representations of stored patterns, learn how to interact with the environment, and build expectations regarding future events. A six year old child has many redundant and plastic connections ready to learn almost anything. After years of learning, the connection density among neurons is reduced, as only the most useful information is retained, and related memories and skills are refined.
Although existing neural network models assume full or almost full connectivity among neurons, the human cerebral cortex is a sparsely connected network of neurons. For example, a neuron projecting through the mossy pathway (of a rat) from the dentate gyrus to subregion CA3 of the hippocampus has been estimated to synapse on 0.0078% of CA3 pyramidal cells (Rolls, 1989). Sparse connections can, at the same time, improve the storage capacity per synapse and reduce the energy consumption of a network.
For the purpose of building intelligent machines, it seems useful to develop a neural network memory that allows the machine to perceive and learn in a manner similar to that of humans. The memory should use a uniform, hierarchical, and sparsely-connected

structure with the capability to self-organize. EI with this type of memory will learn predominantly in an unsupervised manner by responding to stimuli from the environment. The learning process is deliberate, perpetual, and closely related to the machine's goals in the environment.

Having the general purpose of surviving and certain more specific goals, the machine can efficiently organize its resources to process the incoming information and learn the important skills. The creation of goals should result from the machine's interaction with its environment. Therefore, an intelligent machine must have a built-in mechanism to create goals for its behavior and such a mechanism will be called the goal creation system (GCS). The main role of GCS is to develop sensory-motor coordination, goal-oriented learning of perceptions and actions, and to act as stimuli for interaction with the environment. Like the machine's memory, GCS is based on a uniform hierarchical and self-organizing structure. The structure grows in complexity as goal hierarchy evolves. Meanwhile, the goal creation stimulates the growth of the hetero hierarchy representing sensory inputs and a similar hetero hierarchy representing actions and skills.

3.4 Pain signals as motivation

In embodied intelligence research a fundamental question is what motivates a machine to develop into an intelligent, knowledgeable being (Pfeifer & Bongard, 2007). It is an important question since a machine with intelligence is different from a robot that does only the tasks it was designed to do. An intelligent machine must be able to learn and execute various tasks, but the question is what makes it do any of them and in particular what motivates it to strive for excellence in executing these tasks?

To answer this question we may want to ask ourselves what motivates us to get up every morning and go to work. An attempt to formulate an answer to this question was suggested by Csikszentmihalyi (Csikszentmihalyi, 1996), who introduced "flow" theory which states that humans get internal reward for activities that are slightly above their level of development. Stimulated by the "flow" idea, Oudeyer et al. (Oudeyer, et al., 2007) developed an intrinsic motivation system for autonomous development in robots. A robot explores the environment and activates learning when its predictions do not match the observed environmental response. This leads to exploratory learning of the environment and basic sensory-motor coordination. The motivation in such systems comes from the desire to minimize the prediction error and is related to "artificial curiosity" presented by Schmidhuber (Schmidhuber, 1991). A variant of this type of learning was proposed by Barto et al. (Barto et al., 2004). Although artificial curiosity helps to explore the environment, it leads to learning without a specific purpose. It may be compared to the exploratory phase in reinforcement learning - internal reward motivates the machine to perform exploration. It is obvious that exploration is needed in order to learn and model the environment. But is this mechanism the only motivation we need to develop intelligence? Can "flow" ideas explain goal oriented learning? Can we find another more efficient mechanism for learning?

I suggest a goal-driven mechanism to motivate the machine to act, learn, and develop. I suggest that it is the hostility of the environment, as expressed in the definition of EI adopted here, that is the most effective motivational factor for learning. It is the pain we receive that moves us. And it is our intelligence determined to reduce this pain, which responds to the pain and motivates us to act, learn, and develop. The two conditions are

needed together - hostility of the environment and intelligence that learns how to "survive" by reducing the pain signal. Thus pain is good. Without pain there would be no intelligence, and without pain we would not be motivated to develop.
Thus in some strange step in the process of designing a foundation for intelligence, we come back to great philosophers like Plato who stated "if a pain is good, it is because it prevents a greater pain, or leads to a grater pleasure." (Moore, 1993). In philosophy, pain and pleasure are related to motivation for our own actions, as was eloquently stated by Robert Audi (Audi, 2001) - "There are general standards of rationality, including widely held standard of pleasure and pain as generating good prima facie reasons both for action and desire".
The same view that pain is good is shared by medical doctors (Yellon et al., 1996). Brand stated that pain is one of the ways that your body tries to tell you that something is wrong (Brand & Yancey, 1993). Pain can serve as a safety guard and action trigger; for example when exposed to a danger like fire or electric shock, if pain is felt, the body's immediate response is to pull away - a pain action trigger. Many people would die from the infection of a ruptured organ if they did not feel the pain! Leprosy patients lost their body parts not due to leprosy, but their inability to feel any pain at all. Pain also is a great tool for instruction, from the toddler learning to avoid the hot stove, to weight lifters working out and straining muscles, etc. In life, pain serves as a protector against danger or triggers a person to grow spiritually or intellectually after experiencing a cognitive pain.

4. Goal Creation for Embodied Intelligence

In human intelligence, the perception and the actions are intentional processes. They are built, learned and carried out attempting to meet certain goals or needs. Based on primitive needs, people first create simple goals and learn simple actions. Subsequently, by using the learned perception and skills, they build complex perceptions and actions to meet complex goals. It is postulated that this bottom-up process enables a human to find relevant subtasks for a complex task, dividing it into procedures that can be finished step-by-step. The process also generates human needs and these needs or expectations affect human attention to sensory information. In human learning, the rewards are more subjective than objective, and are given by the environment as well as being internally generated.
Pain, as a term for all types of discomforts and pressures, is a common experience to all people. On the most primitive level, people feel discomfort when they are hungry so that they learn to eat and to search for food. They feel pain when they touch burning charcoal so that they learn to stay away from extreme heat. Although, on more abstract levels, individuals experience different motives and higher-level goals, the primitive pains essentially help them to build this complex system of values in order to survive in the environment and to develop skills useful for successful operation.
Neurobiological study facilitated by neuro-imaging techniques, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) etc, supports the suggestion that there are multiple regions of the brain involved in the pain system which form the neomatrix, also called the "pain matrix" (Melzack, 1990). Experiments using fMRI have identified that such a matrix includes a number of cortical structures, the anterior insula, cingulate and dorsolateral prefrontal cortices (Peyron, et al., 2000), and subcortical structures including the amygdala (Derbyshire, et al., 1997) and the thalamus and hypothalamus (Hsieh, et al., 2001).

Two concurrent systems are identified in the pain matrix - the lateral pain system, which processes the physical pains, and the medial pain system, which processes the emotional aspects of pain, including fear, stress, dread and anxiety (Tölle, et al., 1999). The physically harmful stimuli activate neurons in the lateral pain system, and the anticipation of the pains can induce stress and anxiety, which activates the medial pain system. It has also been demonstrated experimentally that the anticipation of a painful stimulus can activate both pain systems (Porro, 2002).

It has been widely accepted for decades that pain has sensory-discriminative, affective, motivational, and evaluative components (Melzack, 1968). The work presented by (Mesulam, 1990) on a neurocognitive network model suggests that the cingulate cortex is the main contributor to a motivational network that interacts with a perceptual network in the posterior parietal cortex. In this work it is proposed that the pain network is responsible for the goal creation process and affects motivation, attention and sensory perception.

In the proposed learning paradigm, the EI machine will use neuronal structures to self-organize the proposed goal creation system (GCS). GCS stimulates the creation of goals on various abstraction levels, starting from the given primitive goals. It is responsible for evaluating actions in relation to EI goals, stimulating learning of useful associations and representations for sensory inputs and motor outputs. It finds the ontology among sensory objects, associates actions and input stimuli, creates needs and affects the agent's attention.

Accordingly, instead of computing a global value system by a typical reinforcement learning (RL) of the embodied machine, the value system is essentially embedded in the hierarchical GCS. In a classical actor-critic RL paradigm, the action is chosen by the action network based on the present sensory (state) input. The critic network evaluates the state-action pair to determine how the action network may improve the selection of actions. However, learning values of state-action pairs in RL is a long and slowly converging process.

Using the GCS, the machine's learning through interaction with its environment becomes an active process since the machine finds the optimum actions according to its internal goals and pain signals. The machine uses internal reinforcement signals, which make learning of state-action pairs' values more efficient. Since internal rewards depend on accomplishing goals set internally by the machine, learning is organized without reinforcement input from the teacher. Once the machine learns how to accomplish lower level goals, it develops a need for sensory inputs required to perform a beneficial action, and this need is used to define higher level goals. Thus the EI agent evaluates and chooses its actions through an integrated system of goals and values that have only loose relations to the primitive goals and external rewards.

In the following sections the concept and structures for the goal creation system will be further developed.

4.1 Goal Creation System

The built-in goal creation and value system triggers learning of intentional representations and associations between the sensory and motor pathways. When the EI machine realizes that a specific action resulted in a desirable effect related to a current goal, it stores the representation of the perceived object involved in such action and learns associations between the representations in the sensory pathway and the active action neurons in the motor pathway. If the produced results are not relevant to the current goal, no intentional learning takes place. Since this usually happens during the exploration stage, such a deliberate learning process protects the machine's memory from overloading with less

important information. This is not to say that a machine cannot learn during the exploratory phase. However, learning in this phase is less intensive and can be based on finding novelty in perceived environment response to EI actions.
Neurons in the goal creation pathway form a hierarchy of pain centers. They receive the pain signals and trigger creation of goals, which represent the needs of the machine and the means to solve its pains. Lower level pains and associated goals are externally stimulated through primitive sensory inputs. Neurons' activation on these inputs may represent a large number of situations that the EI encounters while interacting with the environment. Higher level pains and goals are developed through associations between neuron activities in the sensory-motor pathways that reduce lower level pains. Goals on the lower levels correspond to simple, externally driven objectives, while those on the higher levels correspond to complex objectives that are learned through the machine's actions and are related to finding the best ways to accomplish the lower level goals.

4.2 Fundamental Characteristics of the Goal Creation System

In the proposed goal creation system for intelligent machines, the advancement of EI value and action systems is stimulated by a simple built-in mechanism rooted in dedicated sensory inputs, called "primitive pain". Since the pain signal comes from the hostile environment (including the embodiment of the EI machine), it is inevitable and gradually increases unless the machine figures out how to reduce and avoid it. Pain reduction is desirable while pain increase is not. Thus, the agent has a desire to reduce the pain or equivalently to pursue pleasure/comfort. EI is forced by the "primitive pain" to explore the environment seeking solutions to achieve its goal - reduction of the pain. In this process, the machine will accumulate knowledge about the environment and its own embodiment, and will develop its cognitive skills.
The EI machine may have several primitive pains, and each one of them has its own varying intensity, and requires a distinct solution. At any given time, the machine suffers from the combination of different pains with different intensities. Pains vary over time and the agent needs to set reduction of the strongest pain as its current goal.
We can make references to human learning systems where a similar mechanism is used to induce activity-based exploration and learning. The "primitive pain" inputs for a human include pain, hunger, urge, fear, stress, anxiety, and other types of physical discomfort. The pain usually happens when something is missing. For instance, we feel hungry when we lack the sufficient sugar level in our blood. We feel anxious when we lack enough food or money. We feel fear when we have no protection, etc... This postulation of deficiency in satisfying our goals as a trigger for action and learning makes the proposed goal creation mechanism biologically plausible even at the level of human intelligence. For example, in a newborn baby, a hierarchical goal creation system and value system has not yet been developed. If the baby is exposed to a primitive pain and it suffers, it will not be satisfied until some action can result in the pain reduction. When the pain is reduced, the baby learns to represent objects and actions that helped to lower that pain.
We also need to find and eat food to sustain our activities. A gradually increasing discomfort coming from the low "sugar level" tells us that we must eat. The pain gets stronger and forces us to search for solutions. Similar urges pressure us to go to the bathroom, put on clothes when we feel cold, or not touch a burning coal. The pain warns us against incoming threats, but also forces us to take an action. We feel relief if we take an

action that reduces this pain. Thus pleasure and comfort can be perceived as a reduction of pain and discomfort.

The intensities of the perceived pains prioritize our actions and are responsible for goal creation. For example, the urgent need to go to the bathroom may easily overtake our desire to eat, or even more so to sit through an interesting lecture. In general, the strongest pains will determine the most pressing goals. Thus the pain-based GCS yields a natural goal management scheme.

A primitive pain leads the machine to find a solution and then the solution is set as the primitive goal. Afterwards, the primitive pain will also trigger development of higher level pain centers and create higher level goals. This is based on a fundamental mechanism for the need to act in response to pain and a simple measure for satisfying such a need. I would argue that this simple need to act motivates machine development and may lead to creation of complex goals and means of their implementation. The mechanism of goal creation in a human, and specifically how the human brain controls human behaviors, is not yet fully established in the field of behavioral science or psychology. It is quite likely that the proposed mechanism is different from the way people create their goals. However, it is feasible, simple, and it satisfies our need to establish the goal creation and to formulate the emergence of a goal hierarchy for machine learning. In addition, this goal creation system stimulates the machine to interact with its environment and to develop its skills.

4.3 Basic Unit of GCS

The proposed goal creation mechanism is based on evolving uniform, basic goal creation units. A GCS unit contains three groups of neurons that interact with each other, including the pain center neurons, reinforcement neuro-transmitter neurons and the corresponding connected neurons in the sensory and motor pathways. The basic goal creation unit (GCU) structure is shown in Fig.2. Although as demonstrated in (Starzyk, et al al., 2008), representations of sensory objects or motor actions are best built using distributed groups of neurons in sensory and motor pathways, they are illustrated here as a single neuron for simplicity.

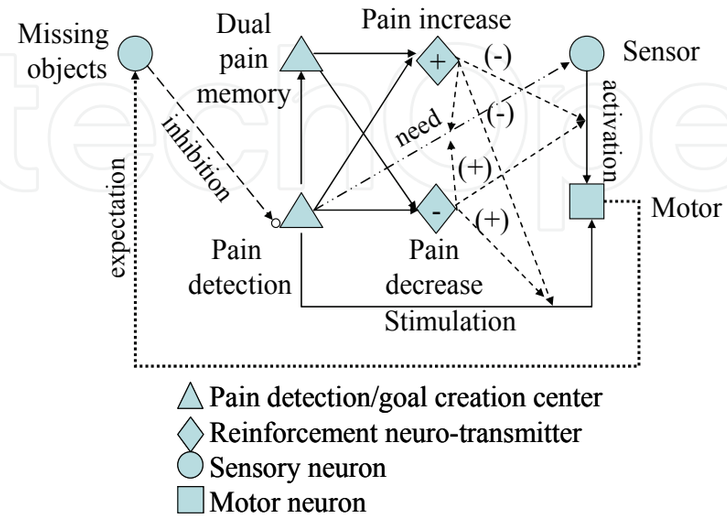


Fig. 2. Basic goal creation unit

The pain detection center is stimulated by the pain signal I_P and represents the negative stimulus, such as pain, discomfort, or displeasure. Since the pain exists due to the absence of certain objects, denoted as "missing objects" in Fig. 2, the perceived object can inhibit the pain signal through the "inhibition" link. Thus the pain detection center is activated by the silence of this sensory neuron. A dual pain memory center stores the delayed pain level, I_{Pd}. The currently detected pain signal and the previous pain signal (in the previously completed event) are compared in the second group, which contains reinforcement neuro-transmitter neurons.

Neuro-transmitter neurons register a decrease or increase in the pain level by comparing signals from the pain detection center and the dual pain memory center. They do not physically connect to any neuron but send positive or negative reinforcement signals to build the associations. The "pain decrease" neuro-transmitter neuron gives a positive reinforcement while the "pain increase" neuron gives a negative reinforcement. The reinforcement signal is calculated in (1).

$$r = I_P - I_{Pd} \tag{1}$$

The third group of neurons in Fig. 2 contains the corresponding active neurons in the sensory and motor pathways that these pain center neurons connect to.

In a GCU, initially, the pain detection center directly stimulates multiple motor neurons. A gradually increasing pain level forces the machine to explore various motor actions by stimulating the motor neurons through initially random connection weights W_{MP}. The machine explores starting from the action with the strongest activation (strongest weights connecting to given pain stimuli). To carry out such action, certain objects which will be involved in this action must be available. Initially, a motor neuron may be associated with multiple sensory neurons by activation weights W_{MS}. The available (active) sensory neurons send activations to the motor neurons so that a certain sensory-motor action pair can be implemented. The direct links from the pain center to the motor neurons force exploration or the implementation of certain motor actions as long as the selected pain persists.

After the action is taken, once the pain reduction or increase is detected by the second group of neurons, a learning signal r is produced to reinforce or weaken the value of an action and the value of the sensory-action pair by strengthening or weakening the stimulation links from the pain detection center to the motor neurons and the activation links from the sensory to the corresponding motor neurons. Pain increase will make the links more inhibitory, while pain decrease will make links more excitatory, as shown in (2)

$$\begin{aligned} W_{MP} &= W_{MP} + r \cdot \beta^n \\ W_{MS} &= W_{MS} + r \cdot \beta^n \end{aligned} \tag{2}$$

where β denotes a smaller than 1 learning rate and n denotes how many times the link has been adjusted.

Meanwhile, since the active sensory neuron representing the object which was involved in the action helps reduce the pain, a "need" link, with weight W_{SP} will be created to connect the active pain detection center to the active sensory neuron using Hebbian Learning. On

the other hand, the object, which was missing and produced the abstract pain signal, becomes available and the neuron representing the object becomes active after the motor action; an "expectation" link with weight W_{SM} will be created to connect the motor neuron and the missing object.
The "need" link and the "expectation" link will be updated by reinforcement learning. The stronger the change in the pain level is, the stronger are the reinforcement signal and the weight adjustment on the involved links. The described interaction of various groups of neurons in the goal creation mechanism and the "stimulation", "activation", "expectation" and "need" links are illustrated in Fig. 2.
This simple mechanism is easy to expand and generalize. In order to generate abstract and complex goals, we will incorporate basic goal creation units into a hierarchy of the goal creation pathway as discussed next.

4.4 Building Goal Hierarchy

A primitive pain is a signal received from the primitive pain sensors. It stimulates the primitive pain detection center. In solving the pain on the primitive level, the machine is stimulated to explore for actions or to exploit the action that relieves the primitive pain. The exploration at first is based on the random stimulation and activation links, or links that were initially (genetically) set to help reduce the primitive pains. Such genetically set links facilitate learning of higher level skills and correspond to built-in skills. Genetic setting of lower level skills may be useful in designing machines that need to develop complex skills.
Genetically set associations between the primitive pain centers and actions also exist in animals. We have genetically wired sensory-motor circuits to sustain basic bodily functions like heart beating, breathing, digestion, etc. A baby cries when it is wet or hungry; it also has well developed skills to eat.
A burning pain from touching a hot plate triggers an automatic pull back reflex. These sensations and actions become gradually associated with the circumstances under which they occurred, leading an intelligent agent to learn basic skills or improve upon them.
To solve the primitive pain from low sugar level, after several random trials, the action "eat", connected with perception of "food", will be rewarded. As a result, the strength of the stimulation link from primitive pain detection center to "eat" and the activation link from "food" to "eat" will be increased. The "need" link is connected from the pain center to "food" and reinforced when such successful action is exploited and rewarded several times. In addition, the "eat" action will trigger expectation of sufficient "sugar level" on the sensory input. Thus, whenever the "low sugar level" pain center sends out pain signals, the "eat" will be excited, prompting the machine for this action.
Since "food" is needed for solving this primitive pain, its absence will lead to anxiety or stress for the machine. A second level pain center representing such stress is created and is called an abstract pain center. An abstract pain center is not stimulated from a physical pain sensor; it only symbolizes a real pain or represents the discomfort of not having the object that can prevent the primitive pain.
When "food" is available and the agent "eats", the primitive pain is relieved. The pain signal disappears and the agent goes back to its normal painless state. As a result, an inhibitory link is developed between the sensory signal representing presence of "food" and the abstract pain center, which means that the existence of the "food" can inhibit the abstract pain.

When "food" is not available, the agent cannot reduce the physical primitive pain. Then, he tries to find a solution to reduce the "abstract pain". Although reduction of the abstract pain does not directly reduce the primitive pain on the lower level, it may be a prerequisite for such reduction.

As specified in the example of a primitive pain center, the agent is forced to explore to reduce the abstract pain. Again, exploration is done based on the initial associations between the abstract pain center and motor actions, and associations between sensory representations and motor actions. The reinforcement neuro-transmitters connected with this abstract pain center update the interconnection weights. Eventually, the reduction in the abstract pain resulting from the action "open" combined with sensory object "refrigerator" indicates that the pain from absence of "food" will be associated with the sensory-motor pair "refrigerator "-"open". It does not matter whether such action (opening refrigerator) was found by pure exploration or by instruction from a teacher. Since once the machine opens the refrigerator, it sees the food and the abstract pain is suppressed, the action will be reinforced. In addition, an expectation link from the motor action "open" to the sensory neuron "food" is built; thus "food" will be expected as the result of the action "open". This expectation link will be used for planning future actions in which a certain action's result can be expected. This process can be illustrated using Fig. 3.

This goal hierarchy can be further expanded vertically. If the agent "opens" the "refrigerator", but the "food" is not found, the machine needs other options to suppress the abstract pain, and subsequently the primitive pain. It may explore the environment by random search, or use instruction. Once it "spends" some "money" (in a store), food is available and then the abstract pain (no food) is reduced. Such action is rewarded and in the future will be more strongly stimulated by the abstract pain center. The "food" is eaten, the primitive pain is suppressed, and the pain signals are reduced. However, when "money" is not available, an abstract pain center on level III is activated with an inhibitory link from "money". Subsequently, the machine needs to learn how to solve the abstract pain on level III related to lack of "money".

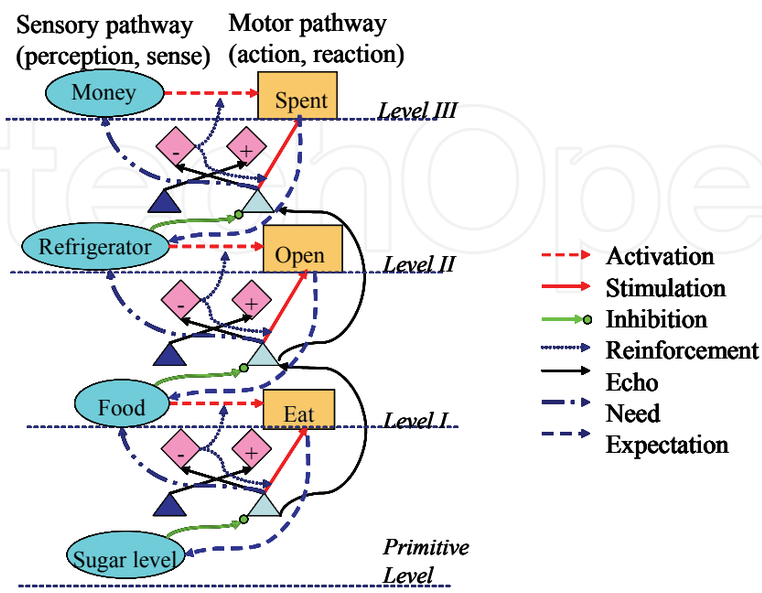


Fig. 3. Creating the abstract pain signals

In the proposed GCS, at every step, the machine finds an action that satisfies its goals, and this action and the involved representations may result in creating further goals. Therefore, in this mechanism, the machine simultaneously learns to associate the goals with deliberate actions, the expected results of actions, the means to represent and obtain objects, and relations among objects. It learns which objects are related to its goals. This helps to establish higher level goals and the means of their implementation. The machine governs execution of actions to satisfy its goals and manages the goal priorities at any given time.

We need both – artificial curiosity and goal creation - to create intelligent systems. We need the first one to explore, and the second one to learn efficiently with a purpose. They complement each other in motivated systems as exploration and exploitation complement each other in reinforcement learning. However, unlike in reinforcement learning, they develop a complex structure of internal goals and rewards that makes learning more efficient, and gives the machine freedom to decide how to approach a given problem. This brings the machine's self-organizing structures a step forward towards general intelligence.

In the proposed model EI uses a hierarchical self-organizing learning memory (HSOLM) for representation building, goal creation, and learning. HSOLM is organized as a hetero hierarchical system of sparsely connected processing units (neurons and their connections). Neurons on different levels of hierarchy handle different tasks. Lower-level neurons are either activated by the sensory neurons and extracted features, or activate motor neurons that control the machine's behaviour. Subsequent level neurons combine the extracted features and represent elements of more complex entities, goals and skills. Information is gathered, associated and abstracted (in an invariant form) as it flows up the memory hierarchy. Top-level neurons represent perceived entities, ideas, goals, skills, and actions.

HSOLM uses three basic pathways – a sensory pathway responsible for perception, a motor pathway responsible for actions, and a goal creation pathway responsible for goal creation, evaluation of actions in relation to goals, learning of useful associations and motivation to act and learn. These three pathways interact on various abstraction levels.

5. The Goal Creation Experiment

The purpose of this experiment is to compare the effectiveness of learning based on the goal creation (GC) system, and reinforcement learning (RL). In order to have a fair comparison, a similar initial neural network topology is used in both learning methods. Networks have similar complexity expressed by the number of input, output and hidden neurons; they also have a similar number of links between neurons and a similar depth of the network. The difference is in how the weight adjustment takes place. While in the RL system, weights are adjusted based only on the external reward/punishment signal, in the GC system, learning is triggered by the local pain centers associated with sensory inputs. In this experiment we assume that both sensory inputs and motor outputs are symbolic, which means that they represent a source (sensory) or action triggered by this output (motor).

The goal creation system built in this experiment is a simplified implementation of the embodied intelligence idea. Thus the system interacts with the environment and is informed about the quality of its actions by an external pain signal. This pain signal is a foundation for setting internal abstract pains and goals to remove these pains.

5.1 Network Organization

Let's assume that the network has to learn how not to go hungry in an environment in which there are limited food resources and advanced skills are required to get them. The machine has six sensory inputs that represent: toys, food, grocery store, bank, office and school. In addition, six motor outputs represent: play, eat, buy, withdraw (money), work, and study. The current state of the environment is determined by the availability of the resources.

A single primitive pain input r(t) that represents hunger is the sugar level in the blood. This primitive pain automatically increases if the machine does not eat for some time. Initially there are plenty of resources around, indicated by a high probability of firing of all the sensory inputs. This allows the machine to learn by exploration how to satisfy its primitive goal (reduce hunger). However, as the machine uses them, the original resources are gradually depleted and need to be replaced. Thus the machine needs to learn how to do so. It will use the GC system to set higher level goals and learn how to accomplish them. Desired sensory-motor pairs and their effect on the environment are listed in the following Table 1. All other sensory-motor pairs are either undesirable or have no effect on the machine's perceived success.

PAIR #	SENSORY	MOTOR	INCREASES	DECREASES
1	Food	Eat	sugar level	food supplies
8	Grocery	Buy	food supplies	money at hand
15	Bank	Withdraw	money at hand	spending limits
22	Office	Work	spending limits	job opportunities
29	School	Study	job opportunities	-

Table 1. Meaningful sensory-motor pairs and their effect on the environment

At each time step the primitive pain increases by 0.1 from its previous level unless it is reset by eating the food. A sensory input probability of various items of category c_i (availability of resource) is described by the function:

$$f_{ci}(k_{ci}) = 1 / [\, 1+(k_c/\tau_c)] \qquad (3)$$

where: τ_c is a scaling factor (resource decline rate), k_c stands for the number of the times a resource was used (initially set to zero). In this experiment τ_c=10. Each time a specific action is taken by the machine (for instance the machine withdraws money from bank account) the corresponding k_c is increased by 1 decreasing the likelihood of this particular resource. This resource can be renewed by invoking a higher level goal after which k_c at a particular level is reset to 0. For instance if grocery was bought, the counter for food was rest to zero and at the same time the counter for money increased by 1.

First, the goal creation experiment is implemented using the reinforcement learning (RL) scheme. The machine must not only learn proper actions, but must adjust its actions to changing environment. Then the same environment is used to test the goal creation system.

5.2 Reinforcement Learning Results

The algorithm of RL used is Actor-critic architecture as shown in Fig. 4. In this scheme, the optimal mapping from states to actions is to be learned during interaction with the external environment based on the primary reinforcement.

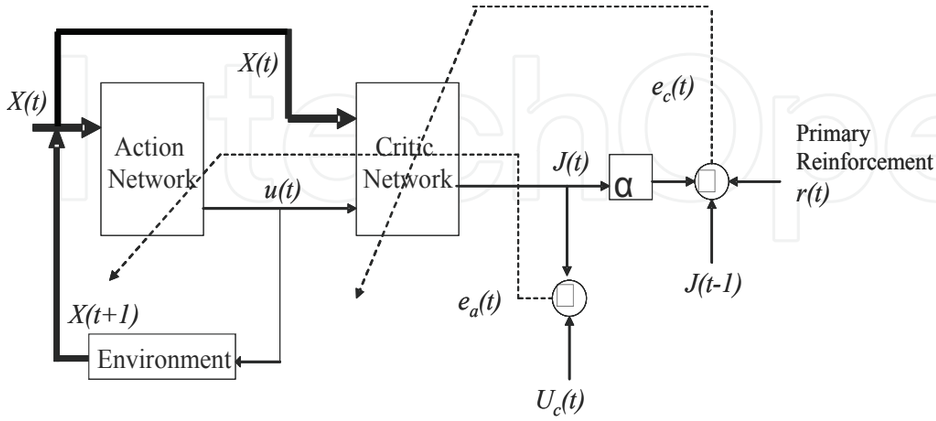


Fig. 4. Actor-critic architecture for RL

The actor-critic architecture contains two components: the action network and the critic network, and both networks are implemented by a multi-layer perceptron (MLP) in this experiment. The action network determines the action u(t) based on current states X(t). The critic network evaluates the state-action value J according to {X(t), u(t)}. The J(X(t), u(t)), also denoted as J(t) value is defined as

$$J(X(t),u(t)) = R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}, \tag{4}$$

where $\gamma, 0 \le \gamma \le 1$, is the discount rate.

The critic network is trained using the temporal difference method (Sutton & Barto, 1998). The critic network directs the action network to produce better action so that J(t) is maximized; thus the action network is trained. A more detailed implementation of this algorithm can be found in (Si & Wang, 2001).

Since the agent has no prior knowledge, the reasonable associations among different objects and actions are not yet built; therefore all possible object- action combinations are considered. The algorithm is described as follows.

Reinforcement learning using AC method:

Step 1). The agent receives input information X(t) and the reward signal r(t) from the environment. X(t) shows the availability of the resources as a binary vector, with "1" indicating being available. The reward signal r(t) is related to the state-action pair taken.

Step 2). The action network (AN) determines the action u(t) from 36 possible action-object combinations based on current input vector X(t). The u(t) is in the form of a binary

vector as well, with "1" indicating the selected action and "0" for all other actions.

Step 3). The critic network (CN) determines the value of this state-action pair J(t).

Step 4). Using the reward signal r(t), CN is trained by TD method. The error function of CN is show in (5).

$$\begin{aligned} e_c(t) &= \gamma J(t) - [J(t-1) - r(t)] \\ E_c(t) &= \frac{1}{2} e_c^{\,2}(t) \end{aligned} \tag{5}$$

Step 5). The weights in CN are updated according to gradient descent,

$$\begin{aligned} w_c(t+1) &= w_c(t) + \Delta w_c(t) \\ \Delta w_c^{(2)}(t) &= l_c(t)\left[-\frac{\partial E_c(t)}{\partial w_c^{(2)}(t)} \right] \\ \Delta w_c^{(1)}(t) &= l_c(t)\left[-\frac{\partial E_c(t)}{\partial w_c^{(1)}(t)} \right] \end{aligned} \tag{6}$$

where $l_c(t)$ is the learning rate of CN.

Step 6). The CN reevaluates the value of J(t).

Step 7). The AN is trained in order to produce action u(t) which has a desired value

$$\begin{aligned} e_a(t) &= J(t) - U_c(t) \\ E_a(t) &= \frac{1}{2} e_a^{\,2}(t) \end{aligned} \tag{7}$$

where $U_c(t)$ is the desired value. The updating algorithm is similar to that in CN and is based on a gradient descent rule,

$$\begin{aligned} \Delta w_a(t) &= l_a(t)\left[-\frac{\partial E_a(t)}{\partial w_a(t)} \right] \\ \frac{\partial E_a(t)}{\partial w_a(t)} &= \frac{\partial E_a(t)}{\partial J(t)} \frac{\partial J(t)}{\partial u(t)} \frac{\partial u(t)}{\partial w_a(t)} \end{aligned} \tag{8}$$

where $l_a(t)$ is the learning rate of AN.

Step 8). Apply the determined action u(t) to the environment, and the environment will give new states X(t+1). Repeat steps 1) to 8).

In a simulated trial, the agent runs for 600 time steps. The pain input (hunger) during this trial is shown in Fig.5.

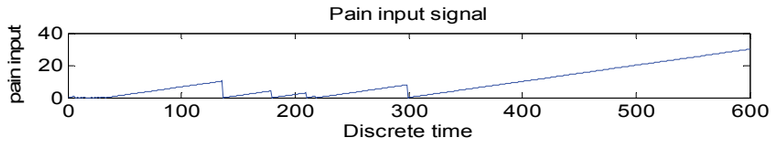


Fig. 5. Results from RL experiment

Initially the RL agent learns how to control the input pain by performing the "eat food" action, but as the food resource is depleted the agent needs to learn another action ("buy food"). The agent learns this action in 103 iterations, and restores the food supply resetting the food counter to 0. After that, it can find food more easily and perform the "eat food" action that reduces the input pain. It repeats the "buy food" action again in iterations number 309 and 407. However, after iteration 299 it never returns to the "eat food" action even though the food is available and the primitive pain increases.

The overall performance by the RL scheme on this experiment can be evaluated from multiple trials. As is evident from the simulation results, it takes a long time for the RL mechanism to adjust to the changing environment and to learn efficient actions. As the system learns, the environment changes requiring higher skills. If the system is not capable of learning these higher level skills in the limited time, it may fail to perform well, since the environment conditions will change and the model initially learned will no longer be satisfactory. In several runs, the changes in the environment happened too quickly for the RL system to keep up with them. In addition to the large number of iterations required in RL, each iteration was more time consuming than the corresponding iteration in the GCS.

5.3 Goal Creation System Results

A goal creation system can be implemented in a general learning scheme similar to the AC method, as shown in Fig 6. In addition to the blocks used in the AC method, it contains a pain network that affects the operation of the critic network, produces abstract pains, and internal reward based on detected inhibition of the abstract pains. These pain detection, goal creation and learning mechanisms were previously illustrated in Fig. 2. The pain network receives sensory input from the environment, including the primitive pain signals r'(t).

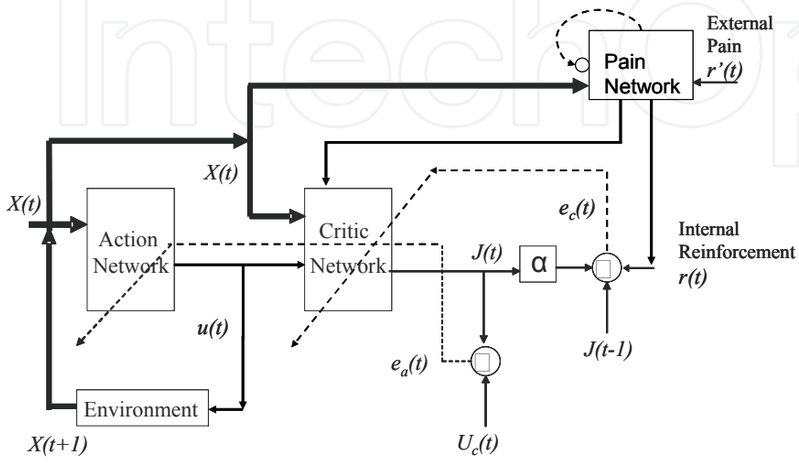


Fig. 6. Goal creation system architecture.

As soon as the machine learns how to satisfy a lower level pain, it identifies the environmental condition that helps it to remove the pain (for instance by supplying food), and creates an abstract pain that activates when these conditions are not met (for instance no food is present). These abstract pain centers have their own bias signals that indicate the importance of this pain compared to other pains. Each time an abstract pain is reduced, it increases the bias weight of the pain center associated with the sensory input that was required to reduce this abstract pain.

A very small decrease is applied simultaneously to all of the bias weights. This small reduction of biases allows the machine to forgo some abstract goals, in the case in which either the environment changed in a significant way and the machine needs to adapt, or the machine learned more effective ways to interact with the environment and replaced former less effective goals with new ones. The machine uses its goal creation approach to learn what to do, and how to adjust to changing environmental conditions. It does so by adjusting pain biases and weights between the pain signals and the critic network.

A typical result of CGS simulation is shown in Fig. 7. This figure shows dynamic changes in the pain signals (including the primitive pain) over several hundred iterations. At first the only pain that the machine responds to is the primitive pain. Once the machine learns that eating food reduces the primitive pain, the lack of food (as observed in the sensory input) becomes an abstract pain. As there is less and less food in the environment, the primitive pain increases again (since the machine cannot get food) and the machine must learn how to get food (buy at the grocery). Once it learns this, a new pain source is created and so on. Notice that the primitive pain is kept under control in spite of changing environmental conditions. On an average run, the machine can learn to develop and solve all abstract pains in this experiment within 200-300 iterations. In all the experiments, the pain threshold was set to 0.1. Below this threshold level no pain is detected.

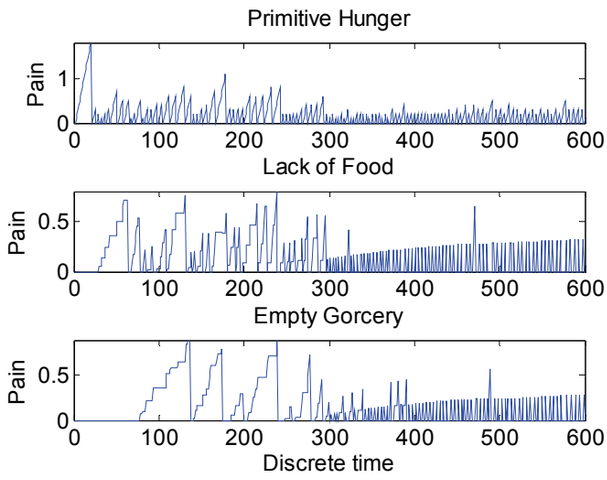


Fig. 7. Pain signals in CGS simulation

Fig. 8 shows a scatter plot that illustrates the selection of a specific action (specified by the sensory-motor pair) in 5 different runs of the GCS. As seen in the attached figure, the machine learns to select useful actions while exploring the environment and selecting the most useful actions repeatedly after it learns a proper interaction scheme.

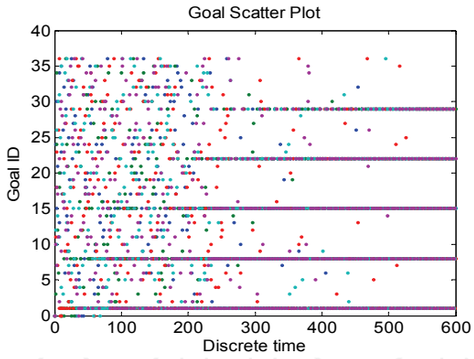


Fig. 8. Action scatters in 5 CGS simulations

The average of all abstract pain signals obtained on the basis of 100 such experiments is shown in Fig. 9. As can be observed, the machine learns to contain all abstract pains and maintain the primitive pain signal at a low level in demanding environmental conditions.

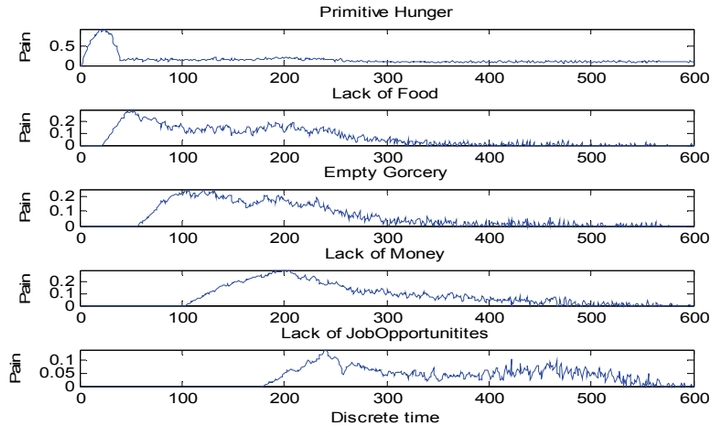


Fig. 9. The average pain signals in 100 CGS simulations

7. Conclusions

This chapter presented a goal creation system that motivates embodied intelligence to learn how to efficiently interact with the environment. The system uses artificial curiosity to explore, and the creation of abstract goals to learn efficiently and purposefully. It develops higher level abstract goals and increases the internal complexity of representations and skills that it stores in its memory. It was demonstrated that this type of system learns better and faster than traditional reinforcement learning systems.

In a striking contrast to classical reinforcement learning, where the reinforcement signals come from the outside environment, GCS generates an internal reward associated with the abstract goal that the machine was able to accomplish. This makes the reinforcement process not observable, and to some degree makes the machine less controllable than one whose operation is based on classical reinforcement learning. The machine's actions are more difficult to understand and explain by an external observer, thus the machine behaves

more like an independent intelligent being rather than a well trained robot.
Instead of the computational-based value system used in typical reinforcement learning, the value system is essentially embedded in the GCS. Initially, the agent acts on very primitive goals and learns the sub-goals or higher level goals through interaction with its environment. At every step, the agent finds an action that satisfies its goals and this action may result in creating further goals. Gradually, the agent learns values of various states for implementing goals. It also learns to associate the primitive goals with its internal states and learns to create higher level goals. At more advanced levels, the agent will be able to understand external instructions and use them as its primitive goals to find the intermediate subgoals or a sequence of abstract goals.
In this scheme, learning that occurs on a higher level of the memory structure to satisfy a higher level goal may have no direct relationship with actions that may satisfy the lower level goals. For instance, an abstract goal of earning money might have been stipulated by a need to buy food or by a need to pay the heating bill or both. In a similar way, accomplishing this higher level goal does not necessarily remove hunger or heat the living space. The only externally observable "primitive pain" may remain unaffected.
This differentiates an externally administered reward related to a "primitive pain" and a reward that satisfies a higher level goal. While we can measure (and thus optimize) the total external reward received to satisfy the primitive goals, it is impossible to measure (and therefore to optimize) the total amount of reward received for higher level goals, as these goals remain mostly unknown to the external observer. This fact differentiates the proposed GCS from the RL scheme, in which the total amount of reward is assumed to be measurable, so that a system can be optimized to maximize the reward. Legg & Hutter demonstrated (Legg & Hutter, 2006) that a mathematically optimal intelligent agent can be built by optimizing the total amount of the rewards that a RL receives. Since the total amount of reward in the proposed GCS cannot be measured from outside, the mathematical proof of optimality presented in (Legg & Hutter, 2006) does not apply. An important question can be formulated as follows – is it an internal reward system and motivation to act, that differentiates an intelligent being from an unintelligent one? I claim that this is true.
In summary, the presented goal creation system motivates a machine to act and develop its cognitive skills in response to externally applied pain signals. It also helps the machine to perceive its environment, learn with a purpose, and respond to changes in environment.

8. References

Agre, P. E. & Chapman, D. (1990). What are plans for?, *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*, Pattie Maes (ed), pp.17-34, MIT Press, Cambridge, MA.

Audi, R. (2001). *The Architecture of Reason: The Structure and Substance of Rationality*, Oxford University Press, ISBN 0-195-14112-1.

Barto, A. Singh, S. & Chentanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills, *Proc. 3rd Int. Conf. Development Learn.*, pp. 112–119, San Diego, CA.

Brand, P. & Yancey, P. (1993). *The Gift of Pain: Why We Hurt and What We Can Do About It.* Zondervan Publishing House, ISBN: 0310221447.

Brooks, R.A., (1991a). Intelligence without representation, *Artificial Intelligence 47*, pp. 139–159.

Brooks, R.A. (1991b). Intelligence without reason, *Proc. 12th Int. Joint Conf. on Artificial Intelligence*, pp. 569-595, Sydney, Australia.

Csikszentmihalyi, M. (1996). *Creativity-Flow and the Psychology of Discovery and Invention*. New York: Harper Perennial, ISBN-10: 0060928204.

Currie, K.W. & Tate, A. (1991). O-Plan: the Open Planning Architecture. *Artificial Intelligence*, 52(1). pp. 49-86, ISSN: 0004-3702.

Dayan, P. & Hinton, G.E. (1993). Feudal Reinforcement Learning, In: *Advances in Neural Information Processing Systems 5*. S. J. Hanson, J. D. Cowan and C. L. Giles (Eds.), Morgan Kaufmann: San Mateo, CA.

Derbyshire, S.W. G., Jones, A. K. P, Gyulai, F., et al. (1997). Pain processing during three levels of noxious stimulation produces differential patterns of central activity, *Pain*, vol. 73, pp. 431–445, ISSN 0304-3959, Elsevier, Amsterdam.

Dietterich, T.G. (2000). Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition, *Journal of Artificial Intelligence Research*, 13, 227-303.

Fu, W.-T. & Anderson, J. R. (2006). Solving the Credit Assignment Problem: Explicit and Implicit Learning with Internal and External State Information. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: LEA.

Hsieh, J.C., Tu, C. H., Chen, F. P. et al. (2001). Activation of the hypothalamus characterizes the acupuncture stimulation at the analgesic point in human: a positron emission tomography study, *Neurosci Lett.* vol. 307, pp. 105–108.

Legg, S. & Hutter, M. (2006). A Formal Measure of Machine Intelligence," *Proc. 15th Annual Machine Learning Conference of Belgium and The Netherlands,* pp.73-80.

Maes P. (1990). Situated agents can have goals, pp. 49-70, *Robotics and Autonomous Systems, 6*.

Maturana, H.R. & Varela, F.J. (1980). Autopoiesis and Cognition - The Realization of the Living: *Boston Studies in the Philosophy of Science* , vol. 42.

McCarthy, J.; Minsky, M. L. & Shannon, C.E. (1955). *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Harvard University, N. Rochester, Aug. 31.

Melzack, R. & Casey, K. L. (1968). Sensory, motivational, and central control determinants of pain, In: Kenshalo DR (Ed). *The Skin Senses.* Pp. 423–439, Springfield: C.C. Thomas.

Melzack, R. (1990). Phantom limbs and the concept of a neuromatrix," *Trends Neurosci*, vol. 13, pp. 88-92.

Mesulam, M.M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory, *Ann Neurol,* vol. 28, pp. 597–613.

Moore, G.E. (1993), *Principia Ethica*. Cambridge: Cambridge University Press, 1903; Revised edition with "Preface to the second edition" and other papers, ed. T. Baldwin, Cambridge: Cambridge University Press.

Moravec, H.P. (1984). Locomotion, vision and intelligence, In: *Robotics Research 1*, M. Brady and R. Paul, eds., MIT Press, Cambridge, MA, pp. 215-224.

Oudeyer, P-Y., Kaplan, F., Hafner, V. (2007). Intrinsic Motivation Systems for Autonomous Mental Development, *IEEE Transactions on Evolutionary Computation*, 11(2).

Parr, R. & Russell S. (1998). Reinforcement Learning with Hierarchies of Machines, In *Advances in Neural Information Processing Systems 10*, MIT Press.

Peyron, R.; Laurent, B. & Garcia-Larrea, L. (2000). Functional imaging of brain responses to pain. A review and meta-analysis," *Neurophysiol Clin*. vol. 30, pp.263–288.

Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence*, MIT Press, Cambridge, MA, 1999.
Pfeifer, R. & Bongard J.C. (2007). *How the Body Shapes the Way We Think: A New View of Intelligence,* The MIT Press (Bradford Books), 2007.
Porro, C.A., Baraldi, P., Pagnoni, G., et al. (2002). Does anticipation of pain affect cortical nociceptive systems? *J. Neurosci,* vol. 22. pp. 3206–3214.
Rolls, E.T. (1989). Functions of neuronal networks in the hippocampus and neocortex in memory, In: *Neural models of plasticity: experimental and theoretical approaches*, pp. 240–265, San Diego: Academic Press, 1989.
Schmidhuber, J. (1991). Curious model-building control systems, *Proceedings Int. Joint Conf. Neural Networks*, pp. 1458–1463, Singapore, vol. 2.
Si, J. & Wang, Y. (2001). On-Line Learning Control by Association and Reinforcement, *IEEE Trans. Neural Networks*, Vol. 12, No. 2, pp. 264-276.
Singh, S.P. (1992). Transfer of learning by composing solutions for elemental sequential tasks. *Machine Learning,* 8, pp. 323-340.
Starzyk, J. A.; Liu, Y. & Vogel, D., (2008). Sparse coding in a hierarchical self-organizing memory with sparse connectivity, submitted to *IEEE Trans. Neural Networks*.
Steels, L. (2003) Intelligence with Representation. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 361, no. 1811, pp. 2381- 2395.
Steels, L. (2004) The Autotelic Principle, in Embodied Artificial Intelligence, *Lecture Notes in AI*, vol. 3139, pp. 231-242.
Steels, L. (2007) The symbol grounding problem is solved, so what's next? In *Symbols, embodiment and meaning*, De Vega, M. and G. Glennberg and G. Graesser, editors, Academic Press, New Haven.
Stewart, J. (1993). Cognition without neurons: Adaptation, learning and memory in the immune system. CC-AL 11, pp. 7-30.
Sutton, R. S. (1984). *Temporal Credit Assignment in Reinforcement Learning.* PhD thesis, University of Massachusetts, Amherst, MA.
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An introduction*. Cambridge, MA: MIT Press.
Tölle, T. R.; Kaufmann, T.; Siessmeier, T. et al. (1999). Region-specific encoding of sensory and affective components of pain in the human brain: a positron emission tomography correlation analysis, *Ann. Neurol.*, vol. 45, pp. 40–47.
Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236): 433-460.
Varela, F. J.; Thompson, E. T. & Rosch, E. (1992). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: The MIT Press.
Walter W.G. (1951). A Machine That Learns, *Scientific American,* 185(2), pp. 60-63.
Walter W.G. (1961). *The Living Brain,* Duckworth, London, 1953, republished by Penguin, Harmondsworth, UK.
Wehner R. (1987). Matched Filters - Neural Models of the External World, *J. Comp. Physiol. A*, vol. 161, pp. 511-531.
Yellon, D.M.; Baxter G.F. & Marber M.S. (1996). Angina reassessed; pain or protector? *Lancet;* vol. 347: pp. 1159–62.



Frontiers in Robotics, Automation and Control

Edited by Alexander Zemliak

ISBN 978-953-7619-17-6
Hard cover, 450 pages
Publisher InTech
Published online 01, October, 2008
Published in print edition October, 2008

This book includes 23 chapters introducing basic research, advanced developments and applications. The book covers topics such us modeling and practical realization of robotic control for different applications, researching of the problems of stability and robustness, automation in algorithm and program developments with application in speech signal processing and linguistic research, system's applied control, computations, and control theory application in mechanics and electronics.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Janusz A. Starzyk (2008). Motivation in Embodied Intelligence, Frontiers in Robotics, Automation and Control, Alexander Zemliak (Ed.), ISBN: 978-953-7619-17-6, InTech, Available from: http://www.intechopen.com/books/frontiers_in_robotics_automation_and_control/motivation_in_embodied_intelligence

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen