

XCR-1: An Experimental Cognitive Robot Based on an Associative Neural Architecture

Pentti O. A. Haikonen

Received: 7 March 2011 / Accepted: 25 April 2011 / Published online: 4 May 2011
© Springer Science+Business Media, LLC 2011

Abstract The experimental cognitive robot XCR-1 is a small three-wheel robot with gripper hands, multiple sensory modalities, and self-talk. The robot XCR-1 is designed for studies and experiments with a new paradigm for cognitive computation, namely an associative neural processing style that inherently and seamlessly combines sub-symbolic and symbolic computation. This operation is realized by using associative neurons and neuron groups organized according to the Haikonen Cognitive Architecture. Recently, there have been many efforts toward machine consciousness, and the Haikonen Cognitive Architecture is one attempt in that direction. Human consciousness is characterized by subjective inner experience that is related to qualia. Accordingly, it can be proposed that true conscious machines should also have some kind of inner experience and qualia. On the other hand, it has been argued that qualia are direct and cannot be artificially realized in symbolic systems. In order to facilitate qualia-related practical investigations, the robot XCR-1 utilizes direct perception processes, with dedicated hardware and without symbolic pre-programmed algorithms. The robot XCR-1 does not utilize microprocessors or programs of any kind. Natural language is one manifestation of symbolic processing. The robot XCR-1 is designed also for experiments with simple speech and the basic grounding of the meaning of words. The experiments with the robot XCR-1 could be greatly enhanced if dedicated associative neuron group chips were available.

Keywords Cognitive robots · Machine consciousness · Cognitive architecture · Associative processing · Robot speech

Introduction

The experimental cognitive robot XCR-1 is a small three-wheel robot with gripper hands, multiple sensory modalities, and self-talk [1]. Experimental robots can be favorably used as test beds for cognitive computation methods, especially those that relate to machine cognition and involve sensorimotor integration with motor action generation and control. The acid test of real-world perception and action provides challenges that are easily overlooked and bypassed in computer simulations.

Usually, experimental robots are built with microprocessors and with a possible link to a master computer, which would execute the actual cognitive computations. These kinds of robots are program controlled, and here, the robot mechanism can be considered as another mechanical output device, not really different from, say, an ink jet printer. More rarely, an experimental robot is realized without the help of a microprocessor or a connected computer. In that case, the robot would not be governed by a program. Associative neural networks can be used to control a robot without any programs or high-level algorithms. The robot XCR-1 is realized in this way.

However, it can be asked why a robot should use some exotic neural hardware instead of the well-tried microprocessors. This question can be put in another way: Considering the robot XCR-1, why less than one thousand transistors are used instead of ten million transistors? Why should one write thousands of lines of computer code when none is needed? Properly designed associative neural

P. O. A. Haikonen (✉)
Department of Philosophy, University of Illinois at Springfield,
One University Plaza, Springfield, IL 62703, USA
e-mail: pentti.haikonen@pp.inet.fi

systems are very efficient, self-learning and require no programming. They are also more close to their biological inspiration, the brain and may, perhaps, be better platforms for potential qualia than symbolic programs.

The robot XCR-1 is one of the few non-trivial robots, which do not utilize microprocessors or digital processing. Instead, it is based on the associative neural architecture of the Haikonen Cognitive Architecture type [2]. This architecture utilizes an associative neural processing style that inherently combines sub-symbolic and symbolic computation. The XCR-1 robot is depicted in Fig. 1.

The Haikonen Cognitive Architecture integrates sensory, memory, and motor systems and functions seamlessly within an associative network. The sensory systems of the XCR-1 extract sensory feature signals that are used to indicate the presence of the sensed real-world features. Signal processing is executed by associative neurons that allow the utilization of the same signals as symbols for completely different entities.

Human conscious experience involves qualia, the subjective qualities of percepts. It can be argued that there can be no human-like consciousness without qualia. Therefore, a conscious robot should also have qualia, but these qualia would not have to be similar to human qualia. The real nature of qualia is not resolved yet, and consequently, the realization of machine qualia is not a straightforward engineering task. The cognitive system of the robot XCR-1 is designed to accommodate the idea that qualia are related to direct representations of sensory information as opposed to symbolic representations that call for interpretation or additional explanatory information [3]. These kinds of indirect symbolic representations would include numeric representations. An example of quale is pain; the feel of pain calls for no interpretation, pain is pain. In a computer, pain could be symbolically represented as a number in a

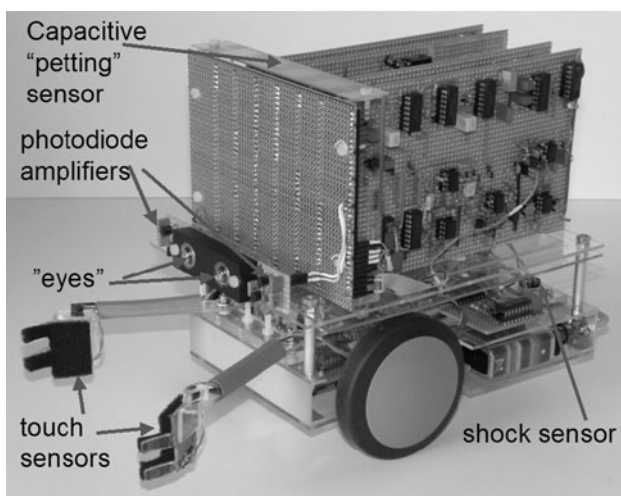


Fig. 1 The XCR-1 robot

file; the higher number, the higher pain. It should be obvious that no subjective experience would be present here. In contrast, in direct systems, pain would appear as a dynamic system reaction with system-wide consequences.

Recently, the value of emotions in cognition has been recognized. Emotional value evaluation may help to decide the order of importance of the perceived threats and affordances. Emotions may also offer templates for fast responses. Without emotions, learning from experience may be inefficient. Thus, a cognitive robot should also incorporate emotions. The cognitive system of the robot XCR-1 is designed for experiments with emotional motivation and behavior control. The body of the robot is shock sensitive, and the robot can be “punished” by hitting it. A “petting” sensor is provided for “reward”. These functions provide the basis for the emotional evaluation of percepts and the modification of behavior according to learned emotional value.

Associative Processing

The robot XCR-1 utilizes associative processing [4], which is based on the use of associative neurons and associative neuron groups. During learning, an associative neuron associates an associative signal vector with the so-called main signal so that later on, the same associative signal vector will evoke the main signal as the output of the neuron. Basically, this operation is similar to the Pavlovian conditioning. Modified Hebbian learning is used; learning takes place when the main signal and the associative signal vector coincide, i.e., are present at the same time. One or more coincidences may be required depending on the actual application.

Associative neurons are grouped into neuron groups allowing the selection of the most strongly evoked signal. This selection is effected by a Winner-Takes-All threshold operation.

Associative neuron groups are also used to associate vectors with vectors. Associative neuron groups are kinds of associative memories. Interference modes that plagued early associative memories are remedied here. Technical details about the used associative neurons and neuron groups are presented in [2]. The principle of the associative neuron group is presented in Fig. 2.

Usually, main signals and associative signals originate from sensory pre-processes and indicate the presence of a given sensory feature. In that case, the associative neuron group can allow the transition from sub-symbolic to symbolic processing; a main signal vector may be taken as a symbol for the corresponding associative signal pattern. For instance, an associative signal pattern may depict a visual object, while the associated main signal pattern,

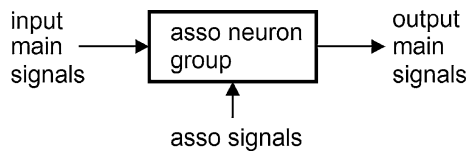


Fig. 2 An associative neuron group associates a main signal vector with an associative signal vector; later on, this associative signal vector will evoke the original main signal vector as the output. The meaning of the input and output main signals is the same

which in itself may depict a sound pattern, may become a name for it. Two cross-coupled associative neuron groups allow associative evocation in both ways: a visual pattern may evoke an auditory name and vice versa.

Architecture

A system architecture describes how the various modules and components of the system are combined in order to achieve the desired overall operation. A cognitive architecture usually combines a number of modules such as sensory modules, memories, reasoning modules, and output modules in various ways that try to produce human-like cognitive functions. A cognitive architecture may be embedded in a computer program, or it can refer to a hardware system.

The robot XCR-1 is an all-hardware system, and its electronically realized neural circuitry utilizes a simplified version of the Haikonen Cognitive Architecture.

The Haikonen Cognitive Architecture is a perceptive system, and its basic building block is the so-called perception/response feedback loop, see Fig. 3. The idea of the perception/response feedback loop has been presented in different forms by various researchers, see for instance Haikonen [2, 5], Hesslow [6], Chella [7].

Figure 3 depicts the general principle of the perception/response feedback loop. A sensor provides sensory information, which is pre-processed into a number of elementary feature signals. Each active feature signal indicates the presence of the corresponding sensory feature. Zero signal indicates that the corresponding feature is not present. Each

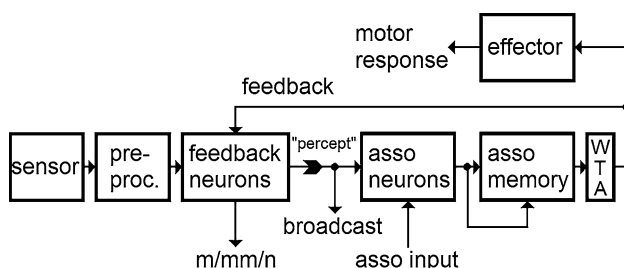


Fig. 3 The principle of the perception/response feedback loop

signal has its own feedback loop; thus, the number of individual signal feedback loops in each sensory module is the same as the number of the extracted feature signals.

A perception/response feedback loop generates a “percept” with the aid of prediction or expectation via the feedback information that re-enters the feedback neuron group. If the feedback vector is close to the sensory input vector, then a match condition is indicated. If they are dissimilar, then a mismatch condition is indicated. This function is useful, for instance, when the robot is searching for some given object. The feedback can also be seen as a priming function; in that case, it can be used to guide attention focus and a selection between a number of percepts.

A perception/response feedback loop acts also as a short-term sensory memory by sustaining the percept signals at a lower intensity level for a while after the disappearance of the actual sensory stimulus.

Each sensory modality has its own perception/response feedback loop modules. The sensory modules are associatively cross-connected. A module “broadcasts” its percepts to the other modules, which receive them via the asso input point. A received broadcast may evoke a sensory signal pattern at the receiving module. This signal pattern is then fed back to the re-entry point of the feedback neuron group and may thus become a new “virtual” percept, one that depicts an internally evoked “imagined” entity.

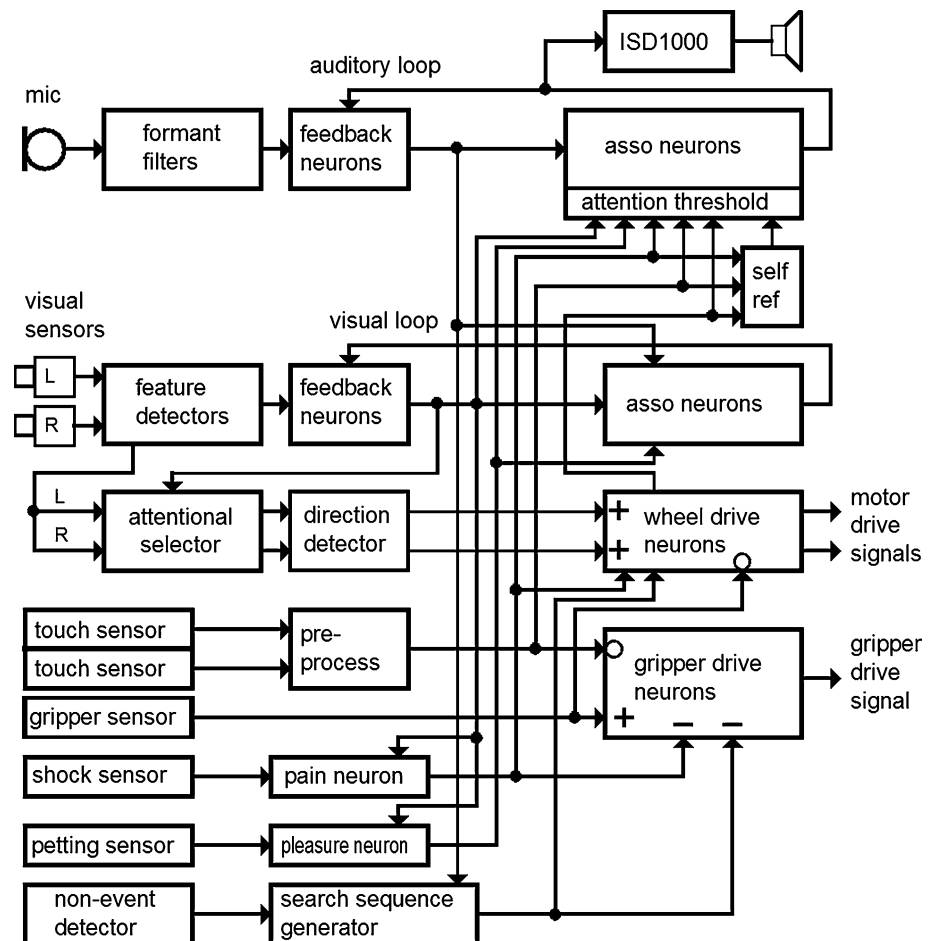
With the inclusion of the “asso memory” block, a perception/response feedback loop will be able to learn and reproduce temporal relationships. The architecture will also operate without this block, but in that case without the sense of past history and time. The realization of this function in the robot XCR-1 is on the waiting list.

The block diagram of the robot XCR-1 is depicted in the Fig. 4. Shown are the auditory and visual perception/response feedback loops, gripper-related sensors, shock sensor, and petting sensor with their connections to gripper drive neurons. Non-event detector is an inner sensor that generates output whenever no other activity is taking place for a while. In this block diagram, the actual motor drive circuits and power conditioning circuits are not shown.

The auditory perception/response feedback loop uses a small electret microphone for the detection of sounds. The sound pre-processing unit consists of simple formant filters that allow the detection of some spoken words. This process is very limited, but suffices for the demonstration of the basic operational principles. The addition of further formant filters and transient detectors would be a straightforward task.

The visual perception/response feedback loop utilizes two large area photodiodes (BPW34) for the detection of active targets. The active targets transmit pulsed infrared radiation and have different pulse frequencies. The visual

Fig. 4 The block diagram of the robot XCR-1, late 2010 status



feature detectors consist of narrow bandwidth filters for the detection of the different targets. Short focal length lenses (from single-use cameras) are used to project the image of a target on the photodiodes. The image of the target projects wholly or only partially on the photodiode depending on the direction of the target, and consequently, due to the parallax effect, the relative amplitudes of the output signals from the photodiodes can be used to determine the direction of the target. This is done by the direction detector, which outputs full direction signal output regardless of the distance of the target. The direction detector provides left-, right-, and straight ahead information to the wheel motor drive neurons, and a closed feedback control loop will arise whenever the robot is approaching a target. This feedback control loop ensures that the robot will move to a position where the object is between the robot's gripper hands.

The gripper system is driven by the gripper sensor that detects objects between the gripper hands. The gripper hands are covered by conductive foam plastic, which functions as the touch sensor; the resistance of the conductive foam depends on the applied pressure. (Black conductive foam

plastic is sometimes used in semiconductor packages; the soft variety is useful.) The gripper motor will stop when the touch pressure reaches a set limit. At that point, the gripper holds the object quite firmly, but not too hard.

A target must be selected when two or more targets are detected at the same time. For this purpose, an attentional selector is provided. This selector is controlled by the visual feedback neuron group output so that the direction detector will always determine the direction of the target that is attended to by the visual perception/response feedback loop. The selector is a very simple inhibiting circuit. Attention in the visual perception/response feedback loop is controlled by emotional value and possible external verbal commands.

It is worth noting that the robot does not utilize any common neural code. The sensory feature signals derive their meanings from the causal connections to the sensed entities. These signals are broadcast to the other modalities, but the receiving modalities do not know the meanings of these signals; they only associate these with their own signals and later on the received signals will be able to evoke the associated signals.

Self-Talk

Little children seem to talk aloud all the time, especially when they are alone. This self-talk is often a running commentary of their instantaneous situation and actions and as such, offers a window into the child's inner mental world.

A similar window is realized in the XCR-1 robot. The self-talk of XCR-1 uses the principles of vertical grounding of the meanings of words [2] and consequently indicates the flow of active concepts in the robot's "mind". The self-talk consists of a limited number of natural language (English) words that are produced by the ISD1000 analog audio memory chip. These audible words are evoked by the actual "inner speech" that is in the form of neural signal patterns in the auditory module.

The auditory signal patterns inside the auditory module are evoked associatively by broadcasts from the other modules. These broadcasts are simultaneous and would lead to overlapping evocations if no limitations were used. Only one word can be uttered at a time; therefore, only one broadcast can be accepted at a time. This limitation can be realized by an attention threshold circuit at the associative input of the auditory module. There are many possibilities for this kind of attentional selection. For instance, emotional value can be used (emotionally most significant broadcasts are accepted first), or a standard word order may be used. Standard word order is easy to implement with minimal hardware and is therefore used here. The attention threshold circuit is made to "scan" the incoming broadcasts and accept only one broadcast at a time. New words cannot be initiated before the completion of the current word; therefore, the scan timing is controlled by the end-of-word information that is provided by the ISD1000 chip. Fall-back timing is used when no words are uttered.

Certain sensory percepts relate to the self-concept, such as movement (it is the subject itself that is moving), touch (it is the subject itself that feels the touch), and pain (it is the subject itself that is in pain). These and similar self-percepts allow and contribute to the emergence of a self-concept. In the robot XCR-1 motion, touch and "pain" percepts are used to evoke a linguistic self-symbol; "me". For instance, the robot may utter: "me hurt", when hit and "me touch", when the touch sensors indicate contact pressure (this utterance may be followed by the name of the touched object). Current wiring does not enable the reports like "me see (something)" or "me hear (something)", but these would be simple additions.

This process may seem trivial, but it is not actually unlike the process in the brain; also there the inner speech is in the form of neural signals, which in turn control the muscle system that produces sound. The actual speech sound-producing system is not important in the cognitive sense.

Self-talk summarizes the activity of all modalities in a symbolic way within the auditory modality.

Emotional System

The first realization of the emotional system of the robot XCR-1 is very simple and is based on the use of "reward" and "punishment" sensory inputs. A "petting sensor" on top of the robot is used to provide "reward" signals. The shock sensor consists of a miniature magnetic earphone (from a toy) that is bolted to the body of the robot. This sensor is mainly sensitive to the vibrations of the body, caused by mechanical shocks at any part of the robot frame. Consequently, the robot can be "punished" by hitting it. The signals from the "petting" and shock sensors are forwarded to "pain" and "pleasure" neurons. These neurons are used to associate positive emotional value, "pleasure" and negative emotional value, "pain" with other percepts.

The general effect of positive emotional value is to favor objects and actions that are associated with "pleasure". Likewise, the general effect of negative emotional value is to avoid objects and actions that are associated with "pain". Consequently, the robot can be trained and motivated to execute certain actions by associating "pleasure" with them. Thereafter, the robot will search for opportunities to execute these actions.

For instance, normally, the robot will approach any visually perceived object. If negative emotional value, "displeasure", is associated with that object, the robot's behavior will change and the object will be avoided. If a neutral-valued and a "pleasure"-valued object are seen at the same time, the "pleasure"-valued object will be favored and approached. The choice between equal-valued objects seems to be a random one.

Auditory word recognition allows the verbal evocation of the "imagined" visual percepts of the test objects, without the actual presence of these objects. This allows the *verbal teaching* of the emotional values of these objects. If the name of an object is spoken and at the same time corporal "reward" or "punishment" is given, then the emotional value of "good" or "bad" will be associated with that object. Later on, when the robot actually observes these objects visually, it will behave according to the associated emotional value of that object; it will avoid bad objects and favor good ones. The robot will also utter the name of the object and its emotional value.

Cross-Associative Information Integration

The verbal teaching experiment demonstrates not only the operation of the emotional system, but also *the*

cross-associative information integration function of the whole architecture. Basically, each module operates on its own, while the overall system operation and behavior emerges from the cross-associative couplings (broadcasts) between the modules. In the verbal teaching experiment, the information integration during learning involves the following more or less parallel steps: the auditory perception of a name of an object; the evocation of an “imagined” visual percept of the object by the perceived name; the perception of the simultaneous corporal reward or punishment; the evocation of the corresponding emotional value; the association of the emotional value with the “imagined” percept of the object; a verbal self-report; and fleeting inner “imagery” evoked by the self-report. The steps during a response are as follows: the visual perception of an object; the evocation of the emotional value of the percept; the corresponding motor response change; a verbal self-report; and fleeting inner “imagery” evoked by the self-report. All these are observable events. These steps are not pre-programmed or pre-wired, but follow from the instantaneous situation and the basic functions of the individual modules.

The philosophical implications of the fleeting multi-modal inner “imagery”, a kind of “instant mental replay”, evoked by the verbal self-reports may not be self-evident at the moment.

Consciousness in the XCR-1

No claims about the consciousness of the robot XCR-1 are made at this point. However, the author has previously proposed that in the Haikonen Cognitive Architecture the difference between conscious and non-conscious operation depends on the global focus of attention [2]. The operation of an individual module is the same, whether the overall operation is considered conscious or non-conscious. The overall operation in the Haikonen architecture is considered to be conscious when all or most modules focus their attention on the same entity or event, because this state would produce certain popular hallmarks of consciousness. In this state, new associative connections can arise that allow the formation of short-term and longer-term memories. This in turn will allow the reporting of the event. This report is not necessarily a report to an outside party; instead, it is a report to the system itself, available instantly and for a while after the incident. These kinds of states occur in the robot XCR-1 and in a simple way would seem to produce at least one popular hallmark of consciousness, namely the reportability, but further studies are required in order to see the actual meaning of this.

Human-style consciousness involves also the inner experience, which is related to qualia. The robot XCR-1 is

designed to have direct perception processes (as opposed to information acquisition and representation by symbols that call for interpretation), and in this way, it should satisfy at least one possible requirement for qualia, namely the directness. Again, further research on the issues of inner experience is needed.

About the Mechanical Platform

The main design criterion for the mechanical platform of the robot XCR-1 was simplicity. However, the mere ability to move around was not considered sufficient. Gripping and touching were considered to be essential additions to the otherwise too trivial set of behaviors.

The robot XCR-1 was designed to have sound perception with one or more microphones. Therefore, it was necessary to minimize mechanical noise in order to avoid auditory interference; thus, the execution of mechanical routines had to be silent. This was achieved by the use of audio quality electric motors without gearboxes. Direct rim drive is used for the wheels; thus, no gearbox is needed. The gripper mechanism is operated by a worm shaft that is coupled directly to the drive motor spindle via a rubber tubing joint.

The silent and smooth operation allows also the acoustical detection of mechanical shocks on the body.

The mechanical construction and the electronic circuits of the robot XCR-1 are designed and built by the author; no commercial robot components are used.

Conclusions

The experimental cognitive robot XCR-1 is a platform for practical experiments with the Haikonen Cognitive Architecture. This architecture utilizes an associative neural processing style that inherently and seamlessly combines sub-symbolic and symbolic computation using associative neurons and associative neuron groups.

A first set of simple circuit modules have been designed for the feasibility study of the approach, and the initial tests look promising. The Haikonen Cognitive Architecture is a working approach, and it can be realized electronically without microprocessors or computers. Several cognitive functions have been realized in minimal forms. Sensorimotor integration including speech production is realized in a direct associative way and as such requires no special interfaces between modules.

Eventual addition of more neurons and synapses would increase the robot’s cognitive capacity, and experiments in that direction are planned. However, currently all neurons and synapses are assembled from general discrete components. This limits heavily the number of neurons that can

be realized within the size and power consumption limitations of the platform. Dedicated associative neuron group-integrated circuits would remedy this problem, but at the moment such chips are not available.

Please see demo videos here: <http://www.youtube.com/user/PenHaiko>.

Acknowledgments I wish to thank Mr. Dylan Drummond for his excellent comments about the text.

References

1. Haikonen PO. An experimental cognitive robot. In: Samsonovich AV, et al., editors. *Biologically inspired cognitive architectures*. Amsterdam: IOS Press; 2010. p. 52–7.
2. Haikonen PO. *Robot brains; Circuits and systems for conscious machines*. UK: Wiley; 2007.
3. Haikonen PO. Qualia and conscious machines. *Int J Mach Conscious*. 2009;2:225–34.
4. Haikonen PO. The role of associative processing in cognitive computing. *Cognit Comput*. 2009;1:42–9.
5. Haikonen PO. An artificial cognitive neural system based on a novel neuron structure and a reentrant modular architecture with implications to machine consciousness. Doctoral thesis. Series B: research reports B4. Helsinki university of technology, applied electronics laboratory; 1999.
6. Hesslow G. Conscious thought as simulation of behaviour and perception. *Trends Cognit Sci*. 2002;6:242–7.
7. Chella A. Perception loop and machine consciousness. *APA Newsl Philos Comput*. 2008;1:7–9.