

# Artificial Perception and Consciousness

H. John Caulfield<sup>a</sup> and John L. Johnson<sup>b</sup>

Fisk University, 10017<sup>th</sup>. Ave., N., Nashville, TN 37208

Commander, V Corps, UNIT 29355, ATTN: AETV-SA (Dr. Johnson), APO AE 09014

## ABSTRACT

Perception has both unconscious and conscious aspects. In all cases, however, what we perceive is a model of reality. By brain construction through evolution, we divide the world into two parts –our body and the outside world. But the process is the same in both cases. We perceive a construct usually governed by sensed data but always involving memory, goals, fears, expectations, etc. As a first step toward Artificial Perception in man-made systems, we examine perception in general here.

## 1. INTRODUCTION

The dictionary meanings of consciousness directly confront the reader with circular definitions all leading back to the human mind, sentience, awareness and self-awareness. When the literature on the nature of consciousness is consulted, each author's conclusions are far afield from those of any other. Worse, they are often if not universally, vague and ambiguous. Yet it is precisely the definition of consciousness itself which we must face as our first step. We must immediately answer the question of how to make a machine think like a person. This necessarily includes resolving the basic problems of self-awareness, consciousness, free will, creativity, language, the meaning of meaning, and intelligence, and others. Our solutions to these problems are presented here.

The outline of the methodology is this: the basic concept of a machine and its interaction with its perceived environment leads to an understanding of the nature of thought and consciousness. Use of known biological mechanisms leads to a resolution of numerous issues including the grounding problem, adaptation, language, and adaptive prediction. The modular nature of these well-defined processes and interactions leads to a single methodology of systematically designing any sort of consciousness in any sort of environment for any sort of entity. These are the fundamentals of conscious intelligence, and their importance cannot be overestimated.

Consciousness will not be redefined here, as it is already defined. From the dictionary:

**Consciousness:** The state of awareness of one's own existence, sensations, and thoughts and of one's environment. Capable of thought, will or perception.

**Aware:** Having knowledge or cognizance. Sentient.

**Sentient:** Capable of feeling. Experiencing sensation or feeling. The human mind.

While the definitions are generally circular, they are the defining criteria. The way to resolve the regressive nature of the definitions is to build on the primary definitions of sentience, the feeling and sensation, the interactions of the world and the body: Sentience is the road to awareness, and awareness is the brick of consciousness. (There are two distinct interpretations of feeling, emotional and physical. Both will be included, in the context explained below.)

There are degrees of consciousness can be approximated as distinct levels but, in practice, may be more of a continuum. These levels are unconscious, sentient, aware, conscious, and self-aware. They are distinguished by the degree to which they model the system and its environment, and by the extent of selective prediction and internal control. A perceptive system would be classified as being at least between the second and third levels.

The system at all but the first level has four primary parts. They are the sensors, the effectors, the rule system by which the sensors and effectors are interrelated, and the external environment in which the system resides. (The environment provides the rules of interaction, and as such is viewed here as a complex rule system that in addition to generating independent stimuli also relates the effectors to the sensors and is thus a “part” of the system.) The unconscious machine’s behavior is open loop, independent of its environment, and it is the only level at which there is no interaction with or consideration of (by the machine) of the environment.

There are four basic algorithmic components. They are a basic reflexive set of rules, an adaptive memory system which incorporates any and all of the numerous adaptive models and whose spatiotemporal associations are regarded simultaneously as rules, maps, and best estimates of the most likely prediction based on previous experience, an innate set of built-in evaluation criteria in the explicit form of additional rules, and an attentional mechanism that incorporates not only competitive interactions among the sensory inputs but also a temporary inhibitory effect that acts only after the winning input channel has been selected (the attentional browser). These four components, the reflexes, the adaptive memory, the judgmental criteria, and the attentional browser, are found in every biological system exhibiting at least awareness or sentience, and are the fundamental building blocks that are used in the design of the system. They will be used repeatedly and in numerous places to generate a given system design.

## 2. Functional Description

Artificial Perception requires a system with the levels of consciousness given above. Figure 1 shows the functional diagram. At the lowest level there is a loop consisting of the sensors, the basic rule system, the effectors, or actuators, and the environment, or the perceived world. The first three are the physical plant, software program, or machine it:

- The sensors include those which detect the external environment (passive sensors), sensors that measure the state and activity of the machine (voltages, fluid levels, temperature), sensors giving

feedback information on the effector devices, and those providing input from other machines and/or the human operator.

- The basic rule system provides a set of fixed reflexive logical rules between the sensors and the effectors. The rules can be simple or complex, fuzzy, crisp, probabilistic, or deterministic, but they are fixed. An analogue is an instinctive, unmodifiable response in a biological entity.
- The effectors are the response actuators by which the machine has an effect on its environment.

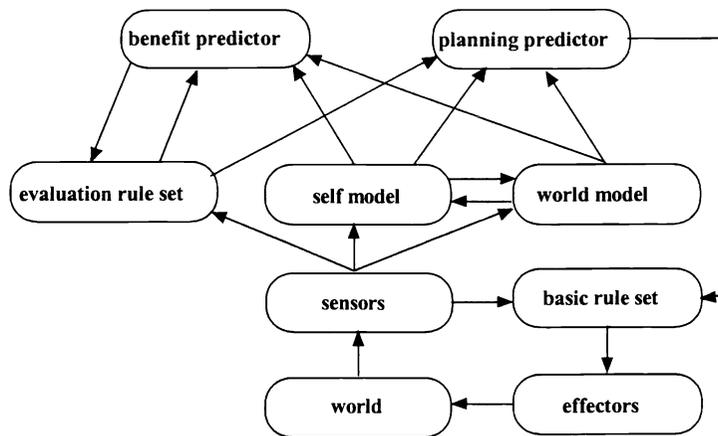


Figure 1. A functional diagram of an Artificial Perception system.

The four basic components are the basic reflexive rule system, an adaptive memory, an attentional browser, and a set of evaluation criteria. The basic rule system has been discussed, above.

- The attentional browser is a type of priority assignment scheme. For example, there are many parallel input and output channels. Wherever there is a conflict as to which subset of inputs to apply, or which reaction to use, an attentional mechanism is applied. There are several types. The simplest is a winner-take-all circuit. A more complex mechanism is a fuzzy controller. Others are the recurrent inhibition used in the adaptive resonance theory (ART), the interaction between the paleocortex and the olfactory bulb, the thalamus, and binocular dominance. These attentional systems are all examples of the general attentional browser, completely automatic and constantly shifting attention from one input or output to the next.
- The basic rule system can be viewed as a spatiotemporal asymmetric associative mapping from input to output. By adding an adaptive memory the system can then manufacture additional maps. The functional diagram labels these new maps according to their type and content, but they are fundamentally the same thing, a map, as the original basic rules.

Unlike the basic rules, they contain information about how the machine interacted with its world. There are two types. First, a sensory input vector is associated with the group of output responses triggered by that vector. The association is causal, and simply restates a subset of the complete basic rule set. It is

significant in that it does not contain anything that was never actually experienced, and represents the current knowledge within the machine of what the machine did. It is a model of the machine. It states its best prediction, based on past performance, of what it will do for a given circumstance. It is the machine's model of itself, the self-model.

The second type causally associates a given effector response with the ensuing sensory inputs. Because this is generally highly dependent upon the external world, this type of associative mapping gives a sample-based model of the external world, the world-model.

Both types are formed in the adaptive memory, and both are repeatedly activated by the current sensory inputs and effector outputs. They then form additional relational maps between themselves, creating by means of standard weight-adaptation algorithms new maps from sensor to effector to sensor and again to effector. This allows something new: The system can generate a prediction of the new sensory state given the original sensory state. Now it can extend its sensory input into the likely futures, and begin to plan ahead. This is the planning predictor.

- But in order to plan, one must have a goal, an objective, and a set of criteria that allow choices to be made. One way (not a very good way) is to choose by means of an attentional mechanism: the reinforcement of the most likely future is done on the basis that its predicted new sensory state best activate the attentional browser. A much better way is found in the human neurological system. In the midbrain there is a collection of nodes, nodules, and elementary organs collectively known as the limbic system, and in particular there are the amygdala, the tegmentum, septum, and the hippocampus which all have been shown to induce physical sensations of pleasure or aversion when stimulated. It is hypothesized that the limbic system provides a fundamental criterion of built-in rewards and punishments, and even that it is a necessary physical element of emotions. What it does do is to generate, not an effector response to a sensor input, but rather a glandular "opinion" of the input, and as such can serve as an elemental set of evaluation criteria. The evaluation, or judgement criteria is thus seen to again be yet another set of rules, but instead of causing a effector response their outputs become additional processing information. They can be viewed in a simple situation as being a variable threshold in a perceptron, for example. They are a set of outputs that measure the value of an input.

With this, the predictive interaction of the model self and model world can be evaluated and the results used to label the most useful likely predicted future.

The evaluation process can continue until an optimum result is predicted, and then the appropriate causal set of responses can be initiated.

This machine has awareness, and is sentient. It is a conscious machine.

The four components, the basic rule set, the adaptive memory, the evaluation criteria set, and the attentional browser are all explicitly buildable using biological neurons and networks with synaptic interconnects. They are each expressible in analytical terms, and have in the past each been encoded in software both as simulations and as processing algorithms. They are not hypothetical conjecture. Our system of Artificial Perception is an algorithm built on established, realistic components.

But it requires more than simply listing the parts and their connections. For example, the transition from a single self-model interaction with a world-model to a lengthy predictive sequence will need attention. The methods of switching from the predictive mode to the effector-sequence mode must be designed. Detailing these aspects will be a significant part of our continuing effort.

- Finally, by extending the functionality to an additional evaluation level, the system can have self awareness. This consists of adding a means of learning by experience what is the most useful evaluation criteria subset, and modifying the value rules themselves. It is similar to the construction of the system's model of its reflexive rules, but here the system builds a model of its evaluation rules. This model can then be used in a predictive sequence to simulate how well a given value rule set will pay off. It predicts the benefit of a set of value rules. The basis of the evaluation of the payoff, itself a value statement, is then in terms of the particular machine capabilities and the specific environment in which it is operating.

There is in the model of Artificial Perception a natural means of generating and grounding a language. It is in the self model, where a passive sensory input from the external world, an event not caused by the machine, can become adaptively associated with a machine response subset. The response gives a new sensor input that is primarily due to the machine's act. This, by the association with the original input, is a valid representation of it, and becomes an icon or word for it whose meaning is grounded in the particular response characteristics of the machine.

This extends to nonhuman systems. For example, the bit string "green" in a digital computer is the icon for a video input intensity state whose green channel is high and whose blue and red channels are low. To the computer, the grounded meaning of the word "green" is that particular video bit state. In a Pavlovian example, the bells and lights become to the animal a word whose meaning is the eminent availability of food.

If a sentient system has the same sensory and effector capabilities as a human, then it will possess a human consciousness. If it has different capabilities, then it will be nonhuman. A sentient, conscious software agent would exist in a world of bits, files, and code, and its sensors and effectors would deal with the objects in that world, and it would generally be a very nonhuman type of consciousness. Likewise, a conscious vehicle would be nonhuman. The issues of how to manage and control them for benefit and not harm will be addressed in the proposed work.

Finally, it is noted that once a system has completed its adaptation/learning and adjusted to its world, it can be described by listing its rules. Accordingly, for very simple systems residing in equally simple environments, the rules can be seen by inspection and adaptation is not necessary. Usually, these are only of academic interest, but can serve as useful illustrations.

### 3. Examples

The first example is a Braitenberg construct. A simple vehicle consisting of two photocells and two motorized wheels exists on an otherwise featureless plane dotted with occasional light bulbs. The bulbs automatically and independently turn on and off for varying lengths of time. The vehicle has a

rechargeable battery, and the charging circuit is activated above a specified threshold of current from the photocells.

The second example is a software agent in the form of a computer game warrior. It consists of three lists. The first list contains the items the agent can sense: the location and actions of the opposing player, the geometry of the local virtual scene, the viability status of the agent, any weapons, etc. The second list contains all the actions allowed to the agent. The third list holds the basic rules initially given to the agent.

## **4. CONCLUSIONS**

We hope we have accomplished several things here, namely:

- Removed some of the unnecessary mystery from the subjects of perception and consciousness
- Shown that constructing an Artificial Perception system is straightforward
- Persuaded you that this is the best way to control ultra complex systems like animals and like many man-made systems and networks.

## **ACKNOWLEDGMENT**

HJC's work was sponsored by The Electric Power Research Institute and by the Army Research Office through contract with Purdue University.