Proceedings of Machine Learning Research 159:63–83, 2021 International Workshop on Self-Supervised Learning

# Artificial Emotions for Rapid Online Explorative Learning

**Paul Robertson** PAUL.ROBERTSON@DOLLABS.COM
*Dynamic Object Langauge Labs Inc., Lexington, MA 02421, USA*

**Editor:** Kristinn R. Thórisson and Paul Robertson

## Abstract

For decades, A.I. has been able to produce impressive results on hard problems, such as games playing in synthetic environments, but have had difficulty in interfacing with the natural world. Recently machine learning has enabled A.I. to interface more robustly with the real world. Statistical methods for speech understanding opened the door to voice-based systems and more recently deep-learning has revolutionized computer vision to the extent that wild speculation now predicts artificial superintelligence surpassing human intelligence, but we are a few major breakthroughs short of that being achieved. We know what some of these breakthroughs need to be. We need to replace supervised learning with unsupervised learning and we need to take on topics like motivation, attention, and emotions. In this article, we describe an architecture that touches on some of these issues drawing inspiration from neuroscience. We describe three aspects of the architecture in this article that address learning through fear and reward and address the focus of attention. These three systems are intimately linked in mammalian brains. We believe that this work represents an attempt to bridge the gap between high order reasoning and base-level support for motivation and learning in robots.

## 1. Keywords:

Online Learning, Attention, Fear, Neuroscience, Alertness, Emotion

## 2. Introduction

As autonomous robots leave their laboratories to work in the real world a new set of needs will become evident. In this paper we describe research on a new formulation of a robot software architecture that draws from nature and provides a framework for integrating perception and actuation, where learning, motivation, and attention play key roles.

Recently, A.I. algorithms have been used to demonstrate impressive results of machine learning (LeCun et al., 2015). Deep Blue beat Garry Kasparov at chess AlphaGo beat the world's best GO player (Silver, 2017); and NASA's ExoMiner has found exo-planets. Many talk about A.I. being able to soon exceed human competence in all domains, but there are some important competence areas that have yet to be mastered.

An autonomous robot today, given sufficient world data can plan a path and conduct a simple mission if all goes well, but expanding it's capabilities to handle the unexpected is not trivial. How should we architect an autonomous robot to deal with an environment that is both passively and actively hostile? How can we build robots that learn from experience in order to be better able to survive in our world?

Robots have made significant progress over the last decade especially when working within a tightly scripted scenario. Robots are not, however, generally aware of what is happening around them outside of their narrowly focused task.

Certain problems have already become recognized as limiting the advance of A.I. and robotics in their course towards superintelligence. **Unsupervised learning** has already been identified as a major missing capability. A related issue is memory. Robots do not build a memory of what they have done outside of their task needs.

Ask your robot what it did yesterday or what it did well and what it did badly. Ask it what it is going to do to improve its performance. The robot is not likely to give good answers to those questions.

The point here, of course, is not that we have not programmed our robots to respond to such questions, it is that we have not seen fit to endow robots with a memory of what they have done, nor to evaluate their accomplishments, even though learning to improve by trial and error. Doing so would lead to more robust robots.

A Boston Dynamics robot can avoid a fall after being kicked by a person, but it doesn't remember being kicked one minute later, nor does it learn to avoid the kick, and certainly not kick back! The kicker, on the other hand, will learn that their attempts to down the robot failed and will learn to kick differently in order to succeed. The kicker and the kickee ought to learn together to ever improve the effectiveness of their kick and their avoidance of the kick much as a GAN does.

A rat will learn to be afraid of situations that cause pain and will attempt to avoid such dangerous situations and will become angry and fight back if cornered. Today's robots do none of those things and will eventually attract abusive humans when humans learn that they do nothing to protect themselves. A robot charged with delivering mail will soon encounter an aggressive dog or school children that would be afraid of the dog and will enjoy harassing a friendly delivery robot.

Even if we decide that robots should not be aggressive, they should surely use their intelligence to avoid situations of danger. They should certainly remember what they have done recently and learned from the experience. If we really want robots to help the elderly and children, they need to understand their world, make sense of what they see and be proactive in helping. Reinforcement learning will eventually learn to not let grandma fall down the stairs, but how many grandmas need to die before that is learned? The fear circuit learns to identify precursors to a potentially catastrophic event and learn to reach to it. Episodic memory is essential for this kind of learning, and can dramatically improve learning for catastrophic cases.

The push to advance the truly hard problems has left a vacuum at the bottom. To achieve superintelligence, we need to fill the holes that we have left along the way. Here are a few more examples: **Generalized Intelligence** that can apply what has been learned in one domain to a new one. **Social Intelligence** that should be able to socialize with other robots and humans in order to work together as teams. **Imitation Intelligence** that can watch someone sliding on a frozen pond, wish to try it itself, imagine doing it itself, try it and fall, want to practice until it can do what the others did, or decide that it is too dangerous and not try again.

### 2.1. What's Missing?

In summary, here is a shortlist of competencies that a super intelligent robot might have: Memory of what has been done and an ability to describe it; a desire to learn and improve itself, in general; the ability to generalize from one domain to another; the ability to work well in teams; the ability to recognize danger and avoid it, and the ability to learn from the world in a totally unsupervised way.

As humans, we value people who work well in teams - **baseball intelligence**. We value people who invent new solutions to old problems - **creative intelligence**, and people who demonstrate **common sense**. Robots who exhibit these kinds of intelligence will be valued.

The previous two paragraphs enumerate some challenges that are only lacking in our motivation to solve them.

Almost everything mentioned above has been demonstrated in some limited sense, and perhaps any of these skills could be learned in specific circumstances and some would argue that if we just enumerated all of the cases and trained up those competencies, we would have succeeded.

The position taken in this article is that to achieve the goals of superintelligence we need generalized and generalizable capabilities that involve comprehension and learning.

Luckily, we have examples of these capabilities in humans and animals. Biology provides solutions that enable animals to survive in a hostile world and while that system is poorly understood today, it provides enough clues to guide an architecture for robotic software that enables better separation of concerns than current architectures and which would ease the problem of scaling up robot performance to handle our complex world.

The approach, described in this article, is biologically inspired and attempts to fill some of the holes while retaining what we have thus far achieved, which is to say, frame an architecture that is inclusive of planners and existing robotic algorithms like SLAM (Thrun and Leonard, 2008) while filling holes that prevent today's robots from being successful as fully autonomous agents that share our world. However hard we try to foresee dangers in our training of robots in laboratory situations, the world has so much complexity that it is hopeless to try to limit learning to the lab. Learning has to continue after the robot is released into the world. We need to architect robots and their learning systems to be continuous, in the real world and exposed to novelty, as part of what it does, every day while performing useful activities.

## 3. Overview of the article

To address some of the gaps alluded to above, we have developed an architecture for Continuous Affective Robot Learning (CARL) and have used it in an initial project where a robot assists a human operator in a repair operation. The robot must understand the video stream as it observed the repair operation. This involves separating the important objects in the scene from the distracters (attention) and to make sense of the sequence of operations performed by the human (sense-making). The robot can describe what is happening and why, using natural language. It can explain what the human should do next, or lead the human back to the correct path if he makes a mistake. The robot can also change its position in order to get a better view of the repair operation, such as when the human occludes the scene. While the robot application had a very limited and practical purpose, it served as the first use case of the CARL architecture. A video of an early demonstration of the system can be found at (Robertson, 2018).

In section 4 we give a sketch of the neuroscience that underlies the CARL architecture. We have taken some liberties in detail in order to provide a clear vision of the inspiration that we have taken in developing the architecture. In section 5 we give an overview of the architecture touching on the key aspects that bridge the gap between biology and robotic architecture. Finally, in section 6 we give an overview of key aspects of the architecture and future directions.

Rather than try to enumerate all the details of the architecture in this article, we limit ourselves, here, to the biological inspiration for the part of the problem that concerns robot safety, learning, and the model for attention.

We use the term "emotion" and "sentiment" in the sense used by (Damásio, 1994). An emotion is a physical response to a situation that is measurable externally, such as by observing muscle contractions involved in facial expressions, or by chemicals released into the bloodstream. A sentiment is a mental state that results from observing that the body is executing an emotional reaction. The sentiments form the basis for higher-level reasoning based on a perception of self. This higher-level reasoning will be discussed in a separate article.

## 4. Neuroscience Overview

For an animal, including a human, and we argue also for robots, when danger presents itself, reacting quickly is essential. Such a situation calls for immediate action, it calls for perception capabilities to be directed at understanding the apparent danger, it may call for problem-solving capabilities, or planning, to be directed as dealing with the emergency, and it represents learning opportunities.

A great deal of what we do as animals and humans that passes as intelligent, can be characterized as learning automatic responses and playing them back when they are appropriate. Reasoning about an intelligent response is necessary too, but we argue that such reasoning should be integrated within a system that is at its core primarily automatic, and learned.

This neuroscience overview follows the basic structure of the CARL architecture. It is not intended to be exhaustive, but rather to motivate the architectural design. We are primarily interested in providing base support for non-supervised learning for which we begin with fear-driven learning in subsection 4.1 followed by reward-driven learning in subsection 4.2. The architecture operates in a perceive-act cycle which mirrors, in essence, the biological cycle which is described in subsection 4.3. The world is cluttered with distractions but to successfully execute useful tasks, a robot must pay attention to what is important while ignoring distracters. Typically today, any notion of attention is hardwired in ad-hoc mechanisms. In our architecture, we wished to expose attention as an architectural feature. The biological inspirations for this are described in subsection 4.4. The last subsection 4.5 describes emotions, which are key to non-supervised learning.

### 4.1. The Fear Circuit

Fear is a great motivator, that is essential in animals for survival and which is implicated in learning and attention. The "fear" circuit, principally centered around the amygdala, provides specialized learning for threat recognition and automated "emotional" responses to the threat. The emotional responses take the form of muscle sequences that may, for example, show fear on the face, the release of hormones to prepare for the danger, a focus of attention on the threat, and a preliminary response to the threat until higher-level reasoning can provide additional guidance. Figure 1 (left) shows a simplified sketch of the fear circuit (Faucher and Tappolet, 2002).

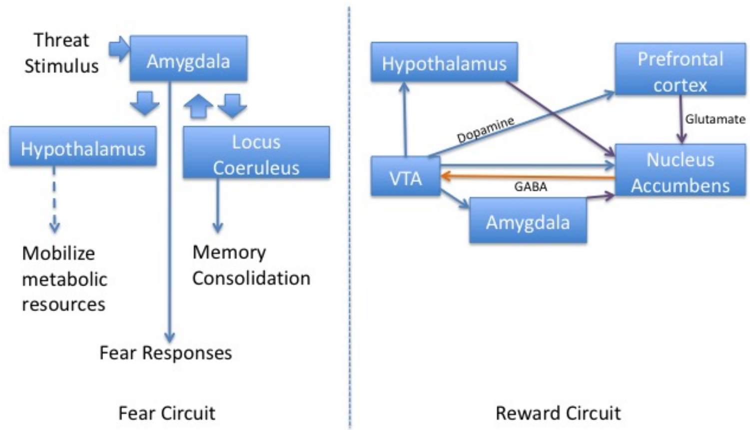


Figure 1: Fear Circuit (left) Reward Circuit (right)

This circuit also guides the learning of new fears in response to exposure to threats and indeed by watching others being exposed to threats. Since exposure to threats can be fatal, it is important to be able to learn about threats by watching others as well as from first hand experience.

When a previously learned threat is detected, such as a spider, a pre-learned fear response causes attention to be given to the threat. This can involve (1) a facial expression of fear, (2) a freeze response to the potential threat, (3) a saccade to the threat in order to gather more information, and the release of hormones, such as *adrenalin* into the bloodstream to prepare to respond to the threat. The freeze response is part of the freeze/flight/fight mechanism. The initial freeze allows for an assessment of the level of danger to be made while preparing to deal with the threat; perhaps the spider is not dangerous. Higher-level reasoning, in the pre-frontal cortex, allows us to better respond to the situation including overriding the fear response in the case that the response is not called for. The released hormones control heart rate, blood pressure, respiration, and digestion and generally prepare the body to react with a huge expenditure of energy, either for flight or for fight. The sensation of the bodily changes in turn is recognized as the *sentiment* of fear by the cortex which in turn focuses attention away from other mental activities.

While certain fears may be innate, such as the fear of spiders and snakes, other fears are learned by the experience of traumatic events. A dog that approaches, growling, showing its teeth, and proceeds to bite is an example of such a traumatic event. The pain of the bite, in addition to triggering the fear response, triggers a learning event. The short period of time preceding the bite can be examined to extract the salient observations leading up to the pain: the approaching animal, the signs of aggression - the visible teeth and the growling sound. These details along with the *context* are learned and will serve as a trigger for the automatic fear response in the future.

Learning to respond to danger is important in order to survive in a hostile world. The fear circuit is thus old in evolutionary terms, but it represents an important form of learning of automated responses and of focusing attention. Dealing with a threat is a high-priority activity that may demand full attention.

### 4.2. The Reward Circuit

Other learned automated responses dealing with survival when not in imminent danger are dealt with by the reward circuit which is similarly old in evolutionary terms and like the fear circuit, described above, involves the recognition of situations in which a pre-learned response would likely yield a reward in the current context. Like the fear response, the reward circuit is built around an emotional response. Figure 1 (right) shows a simplified sketch of the reward circuit (Day and Carelli, 2007; Holmes and Fam, 2013). A key component of the circuit is the neurotransmitter *dopamine*. Dopamine is a neurotransmitter delivered to neurons involved in taking action. Releasing dopamine increases the likelihood of an action being taken. It can be thought of as an *enthusiasm* to act and the sentiment that results from this is *pleasure*. In the context be being hungry, the availability of food makes various actions candidates for obtaining a reward for satiating the hunger. The act of heading towards the candy machine, or pushing on a lever to release food in a laboratory rat experiment, is a candidate action for obtaining the reward. The release of dopamine increases the likelihood of such an action being taken because the neurons involved in taking the action which are already primed to fire based on their relevance to the context are pushed over the threshold by the presence of dopamine. The sensation of pleasure, therefore, allows the reward to be taken in advance of the successful completion of the reward-yielding activity instead of at the completion of the task. If the task is completed and the reward is not obtained, disappointment ensues. As with the pain response, the key here is that between a need and its satisfaction there are actions taken that were key to the successful result, or ultimate disappointment. Scanning over the memory of the period allows for the key actions to be learned as automated responses. In the future, the context of feeling hungry in the presence of a candy machine will trigger the automated response of using the machine to obtain candy. This phenomenon has been very well explored over the decades (Wikipedia, 2013) but for our purposes, what is interesting is the way in which this mechanism provides a mechanism for automated learning. Sequences of actions that lead to a reward can learn, over a sequence of trials, to learn to apportion the reward over the steps in the sequence to enable the selection and replay of action sequences (Keiflin and Janak, 2015). This applies equally well to sequences that were arrived at by trial and error, by accident, or planned. The form of learning about sequences of actions resembles Temporal Difference Learning (TD Learning) (Sutton, 1988; Schultz et al., 1997).

Of similar interest are the ways in which the learning mechanism can be short circuited. It is not simply by trial and error that we learn successful responses. We can also learn by watching others. A person who watches a colleague use the candy machine and obtain candy, can imagine the context of being hungry and satiating that desire by using the candy machine. This learning-by-example (Galef, 1998; Byrne and Russon, 1998; Mataric and Pomplun, 1998; Weber et al., 2000) seeds a learning strategy that will ultimately be tried firsthand. It should be noted that this imitative learning is not yet implemented in the CARL architecture, nor is story based-learning.

A very human extension of this concept is the use of stories to achieve similar learning, both of dangers and of opportunities. Telling a story establishes a context in the mind that resembles a genuine situation. We are able to put ourselves in the position of a character in the story who finds himself in a situation and acts and achieves a successful conclusion or suffers a disappointment. Since the earliest cave drawings for a hunt (Violatti, 2015) to football play diagrams (Moberg, 2004) to fables (Schuster, 2014) to plays and operas (Chong, 2006) the use of language and drawings have allowed transfer learning.

The above-described learning approaches driven by fear and reward can account for a lot of learning as we know it. When higher level reasoning about planning solutions is part of the story and reward allows automatization of successful strategies, we can anticipate learning that approaches human competence without supervision. So far only rather simple learning has been demonstrated, and the imitative and story-based learning are yet to be realized, but we believe that this kind of approach will permit robots to learn in a way similar to the way that we do.

It was Turing (Turing, 1950) who suggested that we should make A.I. learn the way we do:

> *Instead of trying to produce a program to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subject to an appropriate course of education one would obtain the adult brain. Presumably, the child-brain is something like a notebook as one buys from the stationers. Rather little mechanism, and lots of blank sheets. (Turing)*

### 4.3. The Perception Action Cycle

The perception-action cycle is a repeating two-stage cycle of sensory input and responses to those inputs. In biological brains the frequency is determined by the time required to process the incoming sensory input and its sampling is mediated and propagated by alpha waves (Klimesch et al., 2011; Klimesch, 2012). The brain's alpha wave, which is inhibitory, holds off sensory interpretations during the perceived phase cycle and then when the alpha cycle dips, the sensory interpretations are sampled which triggers the act phase which is subsequently propagated across the brain as a traveling wave. This implementation in the brain is fascinating, but of little importance to a robot architecture where the frequency might be driven by the speed of the processing or the speed of the sensors, such as the frame rate of the camera. There is nothing to be gained by mimicking this kind of synchronization in computers – there are easier ways. It does, however show a cycle of perceiving, deciding if to act and if so how, and then act.

The first stage involves the collection and processing of the sensor data. The processing that can be done in one cycle may be limited, requiring multiple cycles for detailed analysis. In addition, in the case of foveated vision, further analysis of a scene may involve eye movements. For the safety of an animal, the speed of certain activities is essential, the recognition of a predator, for example. There is a dual visual pathway that permits fast fear response in humans, but here again, as with many aspects of natural intelligence, it is not clear that such dual pathways offer benefits in a robot architecture.

The second stage may involve performing an automated response, simply updating the current state, nothing at all, or continuing the pre-established goal-directed activity. The cycle also provides for many unsupervised learning activities. Of particular interest are those that relate to the learning of high-speed automated responses that have at their heart, the survival of the organism, in animals, and hopefully in robots too. These systems correspond to "fear" and "reward", described above, is implemented by the fear circuit involving the Amygdala and the reward circuit involving the Accumbens nucleus. Other types of automatic and continuous learning are supported by this cycle which are touched upon below.

### 4.4. Attention

At each perceive-act cycle, there is sensory information from the body as well as from the outside world. The majority of this information does not require attention unless something has changed and even then, we may be better off ignoring them. The system of attention controls what is acted upon and what is ignored (Mangun, 2012; Vossel et al., 2014). Or alternatively, how easily the system is distracted.

The OPS5 system (Martin, 1985) had two modes of conflict resolution (Newell, 1992), named LEX and MEA for lexicographic and means-end-analysis. These corresponded to "easily distracted" and "undistractable". The programmer could choose between one of another which gave rise to two different styles of programming. In reality, a continuum between *focus of attention* and *easily distracted* is necessary, and the setpoint should change based on circumstances. Different sensory modalities can have different levels of impact on attention, and in natural systems, the emotional state of the animal also has an effect on attention.

Focusing on performing an action to the point of not paying attention to something that may be important, may leave the organism vulnerable. Allowing the focus of attention on some action to be distracted by anything makes it difficult to get anything done. Allowing rapid interruptions enough to rule out the need to interrupt the work being done, allows for a compromise. The level of interruption allowable can be mediated by the urgency of what is being done. Anxiety can allow interruptions to interfere with progress, and if the anxiety is warranted the interruptions will have a positive result. Anxiety can be higher if we are in an unfamiliar context.

The ability to modulate focus of attention according to responses to outside stimuli is a very valuable capability. This helps explain the importance of having an architecture with an explicit attention model, and the need for a mechanism, like emotions, to modulate that capability in order to balance goals-directed behavior with self-protection.

### 4.5. Emotions, Sentiments and Learning

Emotions are physical responses to observations that serve to respond to current and future actions that affect the well-being of the organism. Some of these physical manifestations are measurable by visible muscular responses such as facial muscles. Other responses in biological systems include chemical responses that ultimately change mental states, such as the release of adrenaline or dopamine. These in turn affect learning, the focus of attention, and the representations of sentiments that are available as observations that give rise to consciousness of mental state.

An emotion is the embodiment of an automated physical response. The emotion gives rise to an immediate response that may be overridden by higher-level reasoning. The observation of chemical changes in the blood and signature muscular responses, such as facial muscular responses provide *supervision* for learning, and consciousness of sentiments that can drive reflection on the mental state that can lead to reasoned responses. In this way, learned automatic responses can be triggered rapidly in order to respond to the urgency of the situation, while a slower more reasoned response can follow by using higher-level reasoning to override automatic learned responses. An imminent collision must be responded to quickly, but as soon as the collision urgency has been dealt with, a more reasoned trajectory is required. The observation of a potential predator demands an immediate response, such as freezing, followed by an evaluation of the true danger perhaps by overriding the freeze response or perhaps by adopting a fight or flight response.

If a possible aggressor is observed, it is a good strategy to consider the threat real initially, but then to evaluate the level of threat in order to either ignore it or provide a response proportional to the actual threat. A peripheral movement may indicate a threat, but further analysis of the object in question may suggest that it is benign, allowing us to continue with what we were doing.

Other emotions allow attention to be directed toward relevant concerns. In general, an emotion is directed towards an object or objects of a certain type and actions relevant to those objects and the emotion, for example: **Arousal:** Focus attention on object(s) of desire; **Hunger:** Focus attention on food; **Thirst:** Focus attention on drink: **Loneliness:** Focus attention on companionship; **Disorientation:** Focus on orientation; Some of these examples are not appropriate for robots, but maintaining acceptable power levels for the tasks ahead is important as is spatial awareness.

### 4.6. Anxiety

In general, a threat results in a representation of the threat: where and what the threat is, its trajectory, the danger that it poses, and so on, but sometimes the threat cannot be immediately resolved. Imagine that an unexpected sound is heard whose origin cannot immediately be localized and whose cause can equally not be identified. We want to increase our sensitivity to the type of sensory observation in question and to all types of potential dangers in general. We refer to this as anxiety. If I am in a context where I am usually safe, such as at home, and I hear some sounds that indicate the presence of a person or animal that I am not expecting, I will become anxious. I will become more sensitive to similar noises, and my attention will be divided between what I was doing and the heightened awareness of strange noises. The level of anxiety may diminish over time if no further sounds are heard, or may grow if, contrariwise, more strange sounds are heard. At some point, the level of anxiety may become sufficient to make me abandon what I was doing in order to go in search of the suspicious sounds. Anxiety then is an increased sensitivity without a specific object of attention. It is a response that permits increased distractibility.

## 5. The CARL Architecture

The CARL architecture began as an experiment in memory-based learning, in which simple robots learned, by experimentation, simple sequences of actuation to maneuver a maze (Robertson, 2008; Robertson and Laddaga, 2009a). The robots would explore in *learning mode* and the sequences of movements that achieved simple maneuvering goals would be learned for later reuse when asked to navigate the maze to get to a target. By reusing the saved action sequences, the robots could get to the goal without hitting obstacles along the way. The CARL architecture adds in the use of reward and pain as a way of generalizing the means whereby obstacle avoidance and useful action sequences could be learned. CARL also brings the ability to integrate a generalized planner that can plan over the learned action sequences, and a model of attention that supports culprit identification for the learned fear response, and TD learning of action sequences that mimics the ways that these capabilities are implemented in mammalian brains.

The cognitive architectures ACT-R (Fu and Anderson, 2004) and SOAR (Laird et al., 2012), have working memory, episodic memory and long-term memory of action sequences. Both ACT-R and SOAR come from a product system background, where procedural knowledge is represented as symbolic production rules. SOAR and ACT-R both came from the belief that production rules, being fragments of procedural capability, could be learned, eventually, and that in the meantime programs could be written by hand - showing that the production system approach is expressive enough to solve interesting problems.

SOAR and ACT-R while having different characteristics, are both goal-directed and depend upon specific hand-generated goals.

The PolyScheme architecture (Kurup et al., 2011) draws inspiration from Minsky's "Society of Mind" (Minsky, 1988) where the premise is that there are a number of basic mechanisms that form the basis of cognition and that by building support for that small handful of representational underpinnings and providing a mechanism whereby they can interact, we would have an architecture capable of solving a wide range of problems and not just the ones chosen by its designers.

What is lacking in these architectures are the notions of motivation, attention, and self-motivated learning. Obviously, in ACT-R, for example, is it possible to build in the notion of motivation, it is not a core concept.

In SOAR and ACT-R, for example, if there are some rules that are activated (whose preconditions are met), there is something to do. The only question is which rule to select.

The CARL architecture also has working, episodic and long term memories but differs from SOAR and ACT-R in significant ways.

In CARL we have chosen to not make the learned action sequences explicit in symbolic form so as to discourage the idea that CARL is a programming language. CARL comes from the other direction, that by learning rules from experimentation with only very basic drives and motivations, we may be able to demonstrate the emergence of interesting behavior and, one day, intelligent behavior.

Not having a fixed syntax for procedures makes it easier to extend the representation. Our goal was never to handwrite procedures, but rather to learn them. Thanks to architectures, such as those described above, we don't need to prove that complex problems can be solved with these kinds of architectures. Instead, we want to explore learned emergent behaviors that are not programmed in by hand.

CARL has the simple built-in goals to *avoid danger* and to *maintain health*. We can give rewards to aid in simulating experiments with animals. Through numerous parameters that control levels of motivation and activation, we car experiment to find parameter settings that best allow interesting behaviors to emerge. Unlike ACT-R and SOAR, in CARL, it is not a given that there is always something to do, or a problem to be solved. If a charger is nearby the charge action will be available and it will have learned to associate it with a charge reward. It will not necessarily result in the robot charging or doing anything else. if the charge level is high, for example, the motivation may not be high enough for the robot to take the effort to get a charge. There may be nothing that the robot is motivated to do and it will rest dormant until its battery becomes low enough to make getting a charge motivating.

Part of our learning involves a clustering algorithm (Robertson and Georgeon, 2020) that operates in continuous real-time but which requires a slow cleanup process. For this, a sleep cycle is provided. Tiredness raises the bar for the level of motivation necessary in order to act.

Anything that a robot does expends energy and potentially puts it in danger. For CARL doing nothing is the norm. Where CARL differs the most from its predecessors is that it has mechanisms for *motivation*, *attention*, and *simple emotions*.

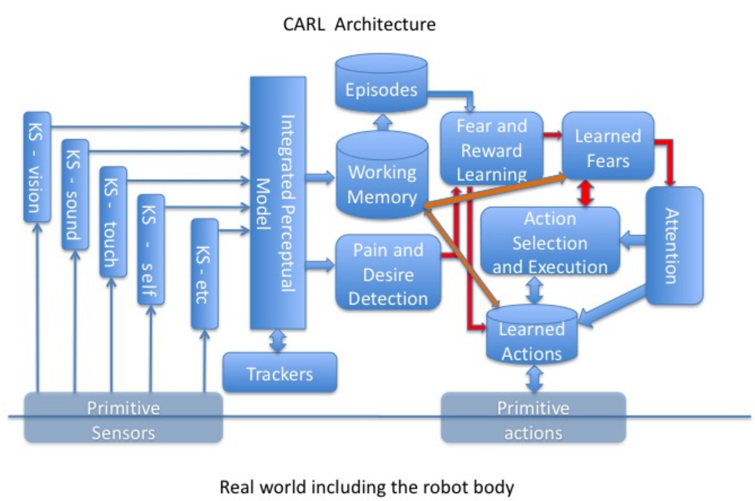


Figure 2: CARL Architecture Diagram

The CARL architecture is shown in Figure 2. The red arrows depict the affective pathways.

Sensors have their own specific modules for interpreting the data from the sensors in order to contribute to an integrated perceptual model. Similarly, trackers provide specialized tracking of the sensed entities from one perception cycle to another. Two outputs are derived from the perceptual model: Updates to the working memory, and pain and desire detection.

"Pain" covers all emergency situations that, for a robot, may include exceeding certain parameters of accelerometers that might indicate free fall or joint angles being outside of desirable limits. For a robot, being knocked over is a likely form of emergency situation, resulting from an encounter with an aggressive or careless child or animal.

"Desire" concerns all non-emergency goals both low level, such as that the battery level remains above 15%, and high-level goals, such as "(go-to location-x)". When the battery is close to the cut off level, getting a charge will be rewarded according to the amount of the charge obtained, so as the battery level runs low, the available reward will be higher and at some point, the "desire" to recharge the battery will overwhelm the attention being given to the foreground task.

### 5.1. Episodic Memory

Episodic memory is key to driving unsupervised learning from both emergency situations and from rewarded situations. In both cases, the pain or reward refers to something that occurred in the recent past. The "Episodes" database contains a parameter-adjustable history of the last "$n$" perception cycles with links back to the raw perceptual data. These episodic entries are used to 1. extract images of the culprit which may be used to learn a fear response to the cause of the emergency, and 2. extract the sequence of actions leading to a reward. An extracted sequence of images of an approaching aggressor can be used to automatically train a CNN recognizer for the aggressor that will be used later as a trigger for a learned fear response. The extracted sequence of actions leading to rewards are stored in the "Learned Actions" memory where the reward assignment and its use in action selection is learned using TD learning.

### 5.2. Learning fear from dangerous situations

When an emergency situation occurs, the robot has several things that it must do, not necessarily sequentially in this order: 1. recover from the emergency, 2. defend against the cause, and 3. learn the cause in order to avoid a recurrence. The activity that caused the problem necessarily occurred in the recent past. The episodic memory exists to allow the recent past to be searched in order to *learn* from the situation. There are a lot of situations that a robot can encounter that lead it into a dangerous situation, some involve encounters with active external agents, others involve the real world and the physical nature of the robot itself. Each time the robot finds itself in such a situation, it represents a learning opportunity to reduce the chances of future reoccurrences.

A battery that has discharged to a dangerous level has undoubtedly dropped below the desired threshold and has failed to be recharged in time to avoid a critical situation. The culprit is the robot itself and the problem is failing to attend to the low battery situation early enough. By searching back through the Episodes, the point at which the search for power was initiated can be found. It was not soon enough. The planner can be invoked to determine how much battery is used in finding a charge, and a fear of running out of power can be established at the minimum level that would obtain a charge before hitting the desired minimum battery level. Establishing the fear will invoke a fear reaction whenever the battery level approaches that point. The fear response will divert attention to the battery and the process of getting a charge will be given priority over the current task ensuring that the robot will seek a charge before getting to a point where it would drop below the desired minimum. Recovery in this case is to plan a route to a charger within the remaining battery power, or to stabilize the robot for a safe shutdown.

If the robot watches an angry dog approach and subsequently attack, and assuming that the dog was correctly identified as the culprit, images of the charging dog can be extracted from the episodes database, that proceeded the attack thereby building a *training set* for learning the fear response. Furthermore, if the learning includes other modalities, such as sound and facial expression, like the showing of teeth, and tail position, the robot can learn to recognize an aggressive dog approaching to attack.

The immediate result of an emotional response provides supervision for the learning of responses depending on the outcome. A successful outcome generates a reward that will reinforce the learning of the successful response, whereas a failure of the automatic response may lead to the future selection of a competing learned response. All of these learned responses are learned within the *context* existing at the time that the emotion is evoked. Thus different learned responses can be learned for different contexts.

Context learning is therefore an important part of what is learned. Context is learned continuously and automatically in CARL through two coupled mechanisms, one involving deep learning (LeCun et al., 2015) and the other involving clustering (**?**). Context learning is a crucial capability of CARL.

"Fear Learning" starts from the "Pain Detection" and entails the determination of the probable culprit in order to extract training data from the episodic memory in the time period leading up to the emergency. From this, learned fear responses are established for rapid detection of possible dangers. The learned fears are checked for continuously, from the integrated perceptual model, and when detected, attention is forcibly directed to the danger, and a learned or innate danger action is triggered.

#### 5.2.1. Diagnosing the Culprit

Diagnosing the cause for an effect is essential for learning, especially to avoid dangerous situations. It is a difficult problem because the approach in TD($\lambda$) learns too slowly for something that could cause death. Natural systems often incorrectly identify causes, but having an irrational fear of something is generally better than not having fear of something that could kill you.

The CARL architecture offers a low-level simple algorithm for cause identification based on temporal and physical proximity, which may be overridden by higher-level understanding. Diagnosis of the cause serves two purposes, the possibility of an immediate physical response, and the learning of a danger. In both cases, higher-level control can override the original diagnosis.

The nature of the sustained pain or wound can guide the diagnosis. Objects close to the pain are more likely to be the cause than ones more distant. The simultaneous emission of a frightening sound by the object, such as a growl, or snarl, increases the likelihood of the object being the culprit. Animate objects are more likely to inflict pain than inanimate objects. Objects that are known to be dangerous are more likely to be the culprit than those that are known to be safe. Unknown objects fall between these two known cases. The diagnosis uses a Bayesian approach with defaults for unknown objects that are either animate or inanimate.

The approach is to use initially seeded probabilities, that are considered to be innately encoded and which can be updated by learning over the life of the robot for probabilities of being potential causes of pain. These include animate versus inanimate objects and distance from the point of pain. Finally, higher-level reasoning can override the low-level diagnosis. If a dog is beside you when you are shot in the leg by a sniper on a rooftop out of sight, the dog may be the first candidate for the cause of pain, but higher-level reasoning can override that diagnosis. This notion of overriding base level reasoning by the prefrontal cortex, in biological brains, is well established. The initial impulse may be to kick the dog, but before that can happen, and certainly before the dog is learned as an object of fear, higher-level reasoning about being shot can prevent harm coming to the dog as well as learning an unnecessary aversion to the dog. Higher-level reasoning will be addressed in a follow-up article.

### 5.3. Learning Reward

A sequence of actions can be considered a finite-state Markov Decision Process (MDP) and in order to attribute the reward and hence learn the expected reward of selecting the sequence we estimate the state value function under a policy $\pi$. Let $\omega^{\pi}$ denote the state value function of the MDP with states $(s_t)_{t\in\mathbb{N}}$, rewards $(r_t)_{t\in\mathbb{N}}$ and discount rate $\gamma$ under the policy $\pi$.

This can be computed incrementally using the TD($\lambda$) algorithm:

If $\omega$ is a vector of weights for the states of the model, the learning achieved after each sequence of actions, $\Delta\omega_t$ is an increment to $\omega$ applied at time $t$.

$$\Delta\omega_t = \alpha(P_{t+1} - P_t)\sum_{k=1}^{t}\lambda^{t-k}\nabla_\omega P_k \tag{1}$$

### 5.4. Perceive Act Cycle

In the CARL architecture, the processing of the observations is performed in two stages corresponding to the alpha cycle: identify, and classify. The identifying step involves trying to identify the cause of the event. It might identify the approaching object as a family pet dog, or as an unknown dog. The sound and motion of a metronome may together identify the source of these two synchronized events as an operating metronome.

The classifying step assigns a level of importance to all identified events. This importance level is a first estimate of whether the robot should act upon the event or simply note it (in working memory). This continuous scale of importance can be loosely described as (1) “safe”, the event can probably be ignored. (2) “dangerous”, the event should be allowed to distract the current activity, and (3) “uncertain”, it may be advisable to pay attention to the event in case it turns out to be dangerous.

The Action Selection and Execution module determines what actions take place during the “Act” part of the perceive/act cycle. During the perceive part of the cycle, the sensors are sampled and the working memory is updated to indicate what has been identified after the noise has been removed and after trackers have integrated the current data into the prior knowledge. For example, the video input runs an object classifier CNN over the latest video frame, it matches up the objects with the prior frame, and for each object, the new position is matched up with its predicted position. The trackers include banks of Kalman filters (Kalman, 1960; Robertson and Laddaga, 2009b) that determine where the object is expected to be. If the object is not close to where it is expected to be, this produces a high salience score for the object. If the object has changed state, for example, every object has several Kalman filters associated with it, each with different system equations, one of which is for free fall. If the object is dropped, the Kalman filter representing free fall in the filter bank will indicate that the object has changed state and that it has been dropped. Similarly, if the object comes to rest or assumes the motion model of one of the hands, this will register that the object has been picked up, or put down. Any change in the state of a tracked object increases its salience score. An object that is relevant to the ongoing activity has a higher salience than an object that is in the background but not participating in the task.

The high-level mission determines the primary objective of the robot. If there is no mission underway, the level of attention is low and any change will have a salience sufficient to be paid attention to. In the course of the mission, certain objects that play a part in the mission will be attended to and the key objects that are expected to change state next in the mission will have the highest attention.

All learned actions that are enabled by the working memory, which is to say that there are bindings for the objects of those actions in the working memory, will be enabled. The actions are rank-ordered based on the attention paid to them and the calculated reward is based on TD Learning will decide which action is chosen to run in the act cycle. A fear response that occurs during the mission will cause the object of the fear to assume high salience by setting the attention to attend to the object of fear instead of what the ongoing task was paying attention to. When that happens, the fear response action will be enabled and will be selected. In this article, we have not described the mission model, but the high-level reasoning part of the system dictates what should be attended to and what objects are expected to change state.

An anxiety level is constantly computed by comparing the context with learned contexts. If the state that we are in falls within a well-defined context that we have learned about, we may be aware of dangers in that context, specific things to watch out for, or the contexts may be considered safe, until such time as we learn otherwise. Being in a known context can heighten awareness of dangers that may be found when in that context. That is increased sensitivity, while in that context to learned fear. If we find ourselves in a context that is unfamiliar, it is here that anxiety plays its role. It is not fear of a particular danger, but rather a sense that we are in unknown territory and that we should allow ourselves to be distracted by anything that might indicate danger.

Contexts are continually calculated and contexts that are familiar and for which fear situations have never been encountered yield a very low anxiety level which enables the level of attention to be high. This results in largely ignoring the background events that are not fear events. Familiar contexts that sometimes have fear events yield a higher level of anxiety, as do unfamiliar contexts. Higher anxiety lowers the level of attention applied, so whereas the attended objects will still rise above the background, a high saliency event in the background, such as an irrelevant object being dropped, or a background object changes its trajectory to be towards the robot, the high saliency will cause it to rise up to a level, where actions associated with those objects will be selectable for execution. Higher anxiety therefore yields greater distraction from the primary task than lower anxiety.

### 5.5. Saliency and Novelty

Saliency is calculated using methods that are sensor-specific and usually work in conjunction with the trackers.

Given an observation consisting of interior and exterior sensory inputs, what is worthy of further attention beyond simply remembering the state? One approach is to measure saliency in terms of local spatial entropy (Gilles, 1998; Kadir and Brady, 2001) or change over time. More generally, model prediction violation (Brown and Friston, 2013). A metronome in the visual scene is salient until its periodicity is modeled after which it ceases to represent a prediction violation, but if the frequency of the metronome suddenly changes, or if it stops, the prediction failure *is* salient. How much attention should be paid to the metronome? That depends upon a number of factors, some learned and some innate: (1) Is the object of attention dangerous? If so, the sudden prediction violation might merit distracting the current task. (2) If the person is anxious, the prediction violation is more likely to be distracting. (3) If the subject has been primed to expect a change in the metronome, it will be treated as important although not strictly distracting, since it is the anticipated event.

A salient event can come from any sensory modality, including: (1) movement detected in the visual field either towards or away from the viewer, (2) a change in color in the visual field, (3) a change in smell, (4) a change in sound, (5) a haptic change, (6) a change in body position, joint positions, and so on. Salient events can occur concurrently in a variety of modalities such as the visual appearance of the dog and the pain of the bite that it is inflicting. Additionally, all events occur in a context that may weigh the relevance of the salient event. Being in a safe location surrounded by a pet, well known to be friendly would assign less importance to a dog that approaches us than a similar dog in an unknown environment.

#### 5.5.1. Focusing Attention:

Control of attention based on context is also essential. When performing a task, we may want to prime for expected outcomes while paying less attention to other distracters. But one wants to respond to danger signs so that work on a task can be temporarily abandoned in order to respond to a threat. If we are already responding to an established threat, we may want to reduce attention to other distracters, such as animals would do for pain, hunger, and thirst. Escaping danger is more important than attending to the pain of a scratch or cut that occurs during flight.

Default attentional responses can be the result of innate values which can be refined by learning. In biological systems, these parameters are often implemented as chemical markers and hormones that affect behavior but in our robot architecture, these parameter values can be represented directly and be observable directly as part of the robot state in the same way that joint angles are. They thus become part of the robot state vector.

Beyond the triggering of learned responses and the associated learning mechanisms associated with the improvement of future responses within contexts, emotions serve as mechanisms for focusing or reducing focus of attention.

To learn that a threat should be learned so as to evoke a fear response in the future, it is necessary to identify the correct culprit. Imagine that I feel a pain in my foot as the result of a dog bite, I should learn that the dog in question is dangerous and thus learn a fear response to that dog. Let's say that prior to the bite, the dog was running toward me and ultimately bit my leg.

The dog, followed by physical contact, followed by a bite, supports identification of the culprit. If there is a picture hanging on the wall when the dog bites me, is the picture the culprit, or is it the dog, or both? We want to be able to learn from a single instance but also from a history of examples. It is reasonable for innate mechanisms to contribute to the diagnosis and for learning by experience to provide for a better diagnosis of culpability.

We know that certain animals can invoke a fear response innately, for example, certain people exhibit an innate fear response to snakes and/or spiders. More generally, animate objects are more likely to bite than inanimate ones like pictures on a wall. In identifying the probability of an object being the culprit, prior, possibly innate probabilities play an important role.

Locality is another useful rule of thumb. When pain is detected, say, on a foot, and upon looking down at the foot we observe a wasp at the site of the pain, and we have prior knowledge that wasps are capable of producing painful strings, we may conclude that the wasp is responsible for the pain especially if the pain is compatible with what is known about wasp stings.

If we stand on a thumbtack and look down to see a wasp on our foot, we may incorrectly associate the pain with a wasp sting. Such culprit identification failures can occur. Similarly, food poisoning can be associated with the wrong food based on temporal locality. These learned fears can be very powerful and be essentially indelible, so misidentification of the culprit is to be proscribed.

A dog that approaches with a wagging tail and a friendly disposition is probably not a danger, whereas an angry looking animal that approached equally rapidly while growling is likely to be a danger. A learned fear response based on a single example is useful because if we are lucky enough to survive one such attack, we may not be so lucky the next time and we want to learn to avoid it the next time.

Outside of innate prejudices, there are many learned things that can help with single-shot learned fear. Animals often exhibit aggression by an emotional; response that involves an angry sound: a growl is a dog, hissing, in some animals, a squawk in a bird, just as humans can produce an angry face. These kinds of responses can be either innate or learned or a combination thereof.

## 6. Conclusions

The CARL architecture builds upon prior experiments involving learning and reusing previously successful sequences of actions. It adds a mechanism for learning the use of these action sequences using TD learning, which resembles reward processing in the mammalian brain and which permits us to perform experiments involving learned behavior with rewards.

The focus on fear and reward-based learning allows for experimentation with unsupervised learning in robots.

CARL provides an architecture in which unsupervised learning can take place during the normal operation of the robot. Emphasis is placed on learning automated responses based on memorizing previously successful responses in equivalent situations and recognizing and learning to avoid dangers.

A, yet to be confirmed, prediction of this kind of automating reuse of learned context-dependent responses, is that it allows the high level planning problems to be easier.

Higher-level reasoning is provided using path planners, generative planners, and temporal planners, while specialized sensors can be plugged in for vision, LIDAR, speech, touch, and so on as knowledge sources (KS). All of whom can contribute to shared working memory.

The attention system provides allocation of computing power as would a traditional scheduler, but control can be taken away to deal with automated responses to threats. The collection of training data for deep learning occurs automatically during the normal operation of the robot and can be processed during a "sleep" cycle.

Unsupervised learning is used throughout the system but in this article, we have chosen to focus on: *Fear*, *Reward*, and *Context*, since they tell a coherent story about how the architecture orchestrates the collection of data for CNN training of fear objects, how TD learning calculates action sequence rewards, and howMDL clustering automatically learns contexts. These three forms of learning are the central features of the architecture. Learning is also used in establishing attentions levels and anxiety levels. In these cases, the intuition, is that if we are doing useful work, a mission, in the case of our testbed, tracking a repair operation, if we are in a familiar setting (context) and there are no fear components in play, we should allow the focus of attention on the key objects to the exclusion of the background irrelevant objects. If the context is less comfortable, (the context is less familiar, or if a potential danger has been detected in the background) we should allow sufficient distraction to avoid an unpleasant surprise. If the robot has a known fearful angry beast heading in its direction, it should forget the mission and execute the flight/fight action!

## References

Harriet Brown and Karl Friston. The functional anatomy of attention: a dcm study. *Frontiers in Human Neuroscience*, 7:784, 2013. ISSN 1662-5161. doi: 10.3389/fnhum.2013.00784. URL https://www.frontiersin.org/article/10.3389/fnhum.2013.00784.

R.W. Byrne and A.E. Russon. Learning by imitation: A hierarchical approach. *J. Behavioral and Brain Sciences*, 6 No. 3, 1998.

Terence Chong. Ethnic identities and cultural capital: An ethnography of chinese opera in singapore. *Identities*, 13(2):283–307, 2006. doi: 10.1080/10702890600698702. URL https://doi.org/10.1080/10702890600698702.

A.R. Damásio. *Descartes' error: emotion, reason, and the human brain*. Quill, 1994. ISBN 9780380726479. URL https://books.google.com/books?id=MocSdi6LXCkC.

Jeremy J. Day and Regina M. Carelli. The nucleus accumbens and pavlovian reward learning. *The Neuroscientist*, 13(2):148–159, 2007. doi: 10.1177/1073858406295854. URL https://doi.org/10.1177/1073858406295854. PMID: 17404375.

L. Faucher and C. Tappolet. Fear and the focus of attention. *Consciousness & Emotion*, 3:2:105–144, 2002.

W. Fu and J.R. Anderson. Extending the computational abilities of the procedural learning mechanism in act-r. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2004.

B. G. Galef. Imitation in animals: history, definition, and interpretation of data from the psychological laboratory. In T. R. Zentall and B. G. Galef, editors, *Social Learning: Psychological and Biological Perspectives*. Lawrence Erlbaum Associates, 1998.

S. Gilles. *Robject Description Learning and Matching of Images*. PhD thesis, University of Oxford, 1998.

Nathan M. Holmes and Justine Fam. How does dopamine release in the nucleus accumbens core relate to encoding of a pavlovian incentive stimulus? *Journal of Neuroscience*, 33 (25):10191–10192, 2013. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.1543-13.2013. URL http://www.jneurosci.org/content/33/25/10191.

Timor Kadir and Michael Brady. Saliency, scale and image description. *International Journal of Computer Vision*, 45(2):83–105, Nov 2001. ISSN 1573-1405. doi: 10.1023/A:1012460413855. URL https://doi.org/10.1023/A:1012460413855.

R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82:34–45, 1960.

R. Keiflin and P.H. Janak. Dopamine prediction errors in reward learning and addiction: from theory to neural circuitry. *j.neuron*, pages 247–263, 2015.

Wolfgang Klimesch. Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12):606 – 617, 2012. ISSN 1364-6613. doi: https://doi.org/10.1016/j.tics.2012.10.007. URL http://www.sciencedirect.com/science/article/pii/S1364661312002434.

Wolfgang Klimesch, Robert Fellinger, and Roman Freunberger. Alpha oscillations and early stages of visual encoding. *Frontiers in Psychology*, 2:118, 2011. ISSN 1664-1078. doi: 10.3389/fpsyg.2011.00118. URL https://www.frontiersin.org/article/10.3389/fpsyg.2011.00118.

U. Kurup, P.G. Bignoli, J.R. Scally, and N.L. Cassimatis. An architectural framework for complex cognition. *Cognitive Systems Researh*, 12:281–292, 2011.

J.E. Laird, K.R. Kindade, S. Mohan, and J. XU. Cognitive robotics using the soar cognitive architecture. In *Cognitive Robotics AAAI Technical Report WS-12-06*, 2012.

Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521:436–444, 2015. URL http://pages.cs.wisc.edu/~dyer/cs540/handouts/deep-learning-nature2015.pdf.

G. R. Mangun, editor. *Neuroscience of Attention: Attentional Control and Selection*. Oxford University Press, 2012. ISBN 0195334361.

Nancy Martin. Programming expert systems in ops5 - an introduction to rule-based programming(1). In *Int. CMG Conference*, 1985.

M.J. Mataric and M. Pomplun. Fixation behavior in observation and imitation of human movement. *Cognitive Brain Research*, 7 No. 2:191–202, 1998.

Marvin Minsky. *Society of mind*. Simon and Schuster, 1988.

B. C. Moberg. Football play scripts: A potential pitfall for federal copyright law? *14 Marq. Sports L. Rev. 525*, 2004. URL http://scholarship.law.marquette.edu/sportslaw/vol14/iss2/9.

A. Newell. *Methaphors for Mind, Theores of Mind: Should the Humanities Mind?* Intellect Ltd, 1992. ISBN ISBN-10: 0893917931.

P. Robertson. Memory-based simultaneous learning of motor, perceptual, and navigation skills. In *Workshop on Machine Learning IROS 2008*, 2008.

P. Robertson and O. Georgeon. Continuous learning of action and state spaces (class). In *Proceedings of Machine Learning Research*, volume 131, pages 15–31, 2020.

P. Robertson and R. Laddaga. A biologically inspired spatial computer that learns to see and act. In *Spatial Computing Workshop, SASO 2009, San Francisco*, 2009a.

P. Robertson and R. Laddaga. Multi-hypothesis self-adaptive network metric tracking. In *Workshop on Machine Learning Meta-Reasoning SASO 2009*, 2009b.

Paul Robertson. Cart: Guiding a user performing a repair task using deep learning, active perception, and plan recognition. 2018. URL https://youtu.be/4dzz3Q1r55s.

W. Schultz, P Dayan, and PR. Montague. A neural substrate of prediction and reward. *Science*, 275(12):1593–1599, 1997.

Joshua Schuster. *The Fable, the Moral, and the Animal: Reconsidering the Fable in Animal Studies with Marianne Moore's Elephants*, pages 137–154. Palgrave Macmillan US, New York, 2014. ISBN 978-1-137-42865-3. doi: 10.1057/9781137428653_8. URL https://doi.org/10.1057/9781137428653_8.

D. et. al. Silver. Mastering the game of go without human knowledge. *Nature*, 550:354–359, 2017.

Richard S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44, Aug 1988. ISSN 1573-0565. doi: 10.1007/BF00115009. URL https://doi.org/10.1007/BF00115009.

Sebastian Thrun and John J. Leonard. *Simultaneous Localization and Mapping*, pages 871–889. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-30301-5. doi: 10.1007/978-3-540-30301-5_38. URL https://doi.org/10.1007/978-3-540-30301-5_38.

A.M. Turing. Computing Machinery and Intelligence. *Mind*, LIX(2236):433–460, oct 1950.

Cristian. Violatti. The meaning of european upper paleolithic rock art. 2015. URL https://www.ancient.eu/article/787/.

Simone Vossel, Joy J. Geng, and Karl J. Friston. Attention, predictions and expectations, and their violation: attentional control in the human brain. *Frontiers in Human Neuroscience*, 8:490, 2014. ISSN 1662-5161. doi: 10.3389/fnhum.2014.00490. URL https://www.frontiersin.org/article/10.3389/fnhum.2014.00490.

S. Weber, M.J. Mataric, and O.C. Jenkins. Experiments in imitation using perceptuo-motor primitives. In *Autonomous Agents*, pages 136–137. ACM Press, 2000. New York.

S. Wikipedia. *Behaviorism: B. F. Skinner, Operant Conditioning, Radical Behaviorism, Behavior Analysis of Child Development, Professional Practice of Behavior Analy*. General Books, 2013. ISBN 9781230634807. URL https://books.google.com/books?id=a5gWnwEACAAJ.