# Artificial Minds with Consciousness and Common sense Aspects

K.R. Shylaja, Dr. Ambedkar Institute of Technology, Bangalore, India

M.V. Vijayakumar, Dr. Ambedkar Institute of Technology, Bangalore, India

E. Vani Prasad, Lakki Reddy Bali Reddy College of Engineering, Mylavaram, India

Darryl N. Davis, University of Hull, Hull, UK

## ABSTRACT

The research work presented in this article investigates and explains the conceptual mechanisms of consciousness and common-sense thinking of animates. These mechanisms are computationally simulated on artificial agents as strategic rules to analyze and compare the performance of agents in critical and dynamic environments. Awareness and attention to specific parameters that affect the performance of agents specify the consciousness level in agents. Common sense is a set of beliefs that are accepted to be true among a group of agents that are engaged in a common purpose, with or without self-experience. The common sense agents are a kind of conscious agents that are given with few common sense assumptions. The so-created environment has attackers with dependency on agents in the survival-food chain. These attackers create a threat mental state in agents that can affect their conscious and common sense behaviors. The agents are built with a multi-layer cognitive architecture COCOCA (Consciousness and Common sense Cognitive Architecture) with five columns and six layers of cognitive processing of each precept of an agent. The conscious agents self-learn strategies for threat management and energy level maintenance. Experimentation conducted in this research work demonstrates animate-level intelligence in their problem-solving capabilities, decision making and reasoning in critical situations.

## KEYWORDS

Artificial Consciousness, Cognition, Cognitive Architecture, Common Sense Behaviour

## INTRODUCTION

Consciousness is a complex mental state that involves the integration of many different mental abilities. Though it is a bold claim that agents can be made fully conscious, agents can only be built with a minimum set of mental abilities that can make them conscious. The main purpose of this research is to understand, adopt, and test some of the principles and complexities of animate consciousness and common sense on either robots or synthetic agents. This paper aims at proposing a self-configurable computational model for implementing and testing animates consciousness and common sense critics using a cognitive approach.

## BACKGROUND

There are many existing cognitive architectures that are built to test and implement cognitive capabilities of the human mind. The Emotion Machine Architecture (EM-ONE) demonstrated human common sense thinking capability in the Roboverse environment (Singh, 2005; Minsky, 2006). The Computational Model for Affect Motivation and Learning (CAMAL) (Darryl & Suzanne, 2004; Darryl, 2010, 2002, 2001) architecture emulates emotions. The Society of Mind Cognitive Architecture (SMCA) investigated the concept of mind as a control system by using the "Society of Agents" metaphor that uses fungus eater testbed (Vijaykumar & Darryl, 2008; Vijaykumar, 2008). The CERA-CREMIUM architecture of Arrabales (2009) demonstrated different levels of consciousness on artificial agents. The research work presented in this article attempts to address the problem by using ideas from AI and cognitive science. Cognitive capabilities of animals and humans are evident when they exhibit abilities such as learning, remembering, perceiving, thinking, decision-making, recognizing, and visual, verbal, and language skills in their usual interactions. Cognitive science proposes theories to build artificial minds based on natural mind architectures called cognitive architectures (Anderson, 1993; 1996; Armstrong, 1968). These architectures help in modelling a range of human behaviors into machines to make them intelligent across a diverse set of tasks and domains. The main focus of any cognitive architecture is to represent, organize, utilize, and acquire the knowledge while performing the task (Newell, 1972;1990;1992).

### Theory of Conscious Agents

According to Russell (2003), an agent is "anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators." The mapping between the percept sequence and the action chosen is called the agent function, whereas the internal processes that choose actions according to the percept sequence are the agent programs.

Most of the human mental processes are unconscious though humans are considered as highly conscious agents (Bargh & Morsella, 2008). The conscious agents are the entities that exhibit intelligent behavior with properties such as autonomy, reactiveness, and pro-activeness or being rational. According to Donald D Hoffman (2014), the mathematical definition of a conscious agent involves three mental processes such as perception, decision making, and action. An agent being in a conscious state can also have subjective experiences, wishes, beliefs, desires, and complex thoughts (Block,1995; 2002; 2002; 2007; Shoemaker,1996). It should be able to understand a relatively complex sequence of actions at an abstract level and respond to such situations (Franklin, 2009). A minimum prerequisite for conscious agents is social interaction with its peers in the environment.

## Conscious Agents with Common sense Critics

Common sense is a set of beliefs or propositions that are considered true by most people as they experience the same and by virtue of this would be the obvious true judgment. The other definition of common sense is the "sense of things given by each sensory organ which can be interpreted and integrated into a single impression among multiple possible impressions (Antonio & Giuseppe, 1999; Barry Smith & D.W. Smith, 1995). " These common sense responses in a person can be presented in either a conscious or an unconscious state. The quality of actions resulting from common sense drastically differs from actions owing to conscious thinking or under the influence of emotions. Common sense is sometimes accepted among people belonging to the same place, culture, and occupation. In some situations, actions attributed to common sense can improve performance and in some cases, they may not even trigger any reaction. In this research, common sense is considered to trigger when agents are not influenced by emotions.

The work presented in this article focuses on building intelligent agents that are highly conscious of external world and adopt common sense and consciousness strategies to respond in dynamic environments.

## Theories of Consciousness

Dennett's (1991) Multi-Draft-Model (MDM) and Bernard Baar's (1997) Global Workspace Theory (GWT) suggests that the human brain is a parallel set of specialized unconscious networks of information processors. Each of these processor networks functions independently and becomes conscious based on the context. The information sensed is broadcasted to all these networks by a context processor. These networks cooperatively work together to produce a cognitive task by using a central information exchange memory area called "Global Workspace." This memory is limited in size and is a short-term resource. The sensed information is laid out by each of the processor networks to understand different perspectives of the information. These processors then send the processed view to other processors by getting an access to global workspace and make it a conscious experience. As it is a short-term memory and a limited resource, each conscious experience can only stay for a few seconds and then

switch to the next experience. These mechanisms enable GWT to account for the ability of consciousness to handle novel situations, its serial procession of states, and the transition of information between consciousness and unconsciousness (Baars and Katherine McGovern, 1988, Baars,1998,2003; Baars, Franklin S, 2003).

According to the BDI model proposed by Bratman (1988), the practical reasoning process of humans has two steps: (a) consider all the desires of an agent and (b) select the most desirable one by mapping it to its current belief set. In this deliberated step, the agent pursues and adopts an intention to achieve a desire. The intentions are persistent in nature and recur till they are achieved. If the intention chosen fails repeatedly to achieve the desired state, the agent can drop this and update its belief set. Hence, intentions are the prime reason for an agent to change its future belief set. In each intentional state an agent considers or adopts options that are consistent with that intention. In principle, intentions justify the possibility of achieving a goal state in the current state. The second step in practical reasoning process involves generating a plan of actions based on goals, beliefs, and actions of agents by using means-ends reasoning. A running agent adopts varying plans that are triggered by external or internal events. This plan involves a sequence of actions that are selected based on the available set of beliefs.

As discussed above, practical reasoning systems are designed based on BDI models that help in achieving the goals of agents. Hence BDI models and GWT principles are adopted in building Consciousness and Common sense Cognitive Architecture (COCOCA) architecture.

## ANIMATE TESTBED SETUP

The animate testbed has been used as a computational tool to measure the performance of ecologically inspired agents. These agents consciously survive in an environment by demonstrating natural behavior of survival in unknown environments.

The Testbed setup considered for COCOCA agent is a 30X30 grid environment, which can also be varied to different sizes. This is filled with food parameters, trap points, and obstacles, as shown in Figure 1. The raw fruits, dry fruits, and juicy fruits are created with different calorific values. All parameters are randomly created in random positions. Both the agent and the attacker consume one unit of energy for each move in normal grids and 2 units in trap grids. These trap grids are hidden and are perceived as normal grids by agents. Both the attacker and the agent are created with 100 units. The agents and attackers always maintain energy for at least 30 next moves without food.

The attackers always look for a nearby agent as a prey to satisfy their food requirement. The attackers ensure that they do not attack agents when in group and agents with common sense use this knowledge to escape the threat. A safe-zone is also created in an environment where the attacker does not enter and this is known to the common sense agents too.

The agents of different levels of consciousness have been created to test their performance in a given scenario. The agents' most preferable food is the raw fruit,

which gives them maximum energy. An agent collects dry fruits as energy store for long-term survival. Here, the resource parameter is the dry fruit and the agent always tries to collect maximum number of dry fruits. The juicy fruit is the least preferred food and has less caloric value. The consciousness of the agents has been tested by measuring their survival rates in the presence of an attacker. The other performance parameters are the number of dry fruits collected and the energy level maintained.

The agents simulated in animate testbed can belong to one of these categories: FSM agent, threat-conscious agent, energy-conscious agent, conscious2 agent, learning agent, and a common sense agent. Each agent type simulates different types of cognitive capabilities that contribute to the agent's consciousness (see Table 1).

## THE PROPOSED AND DEVELOPED COCOCA (CONSCIOUSNESS AND COMMON SENSE COGNITIVE ARCHITECTURE)

The COCOCA is a six-layer cognitive architecture developed for building control systems for agents that are conscious to survive in new and dynamic environments. The layers included are reflexive, reactive, deliberative, conscious, meta-reasoning, and common sense, as shown in Figure 2. The agents of each layer exhibit different levels of intelligent behavior in the domain-specific tasks as they have varying cognitive
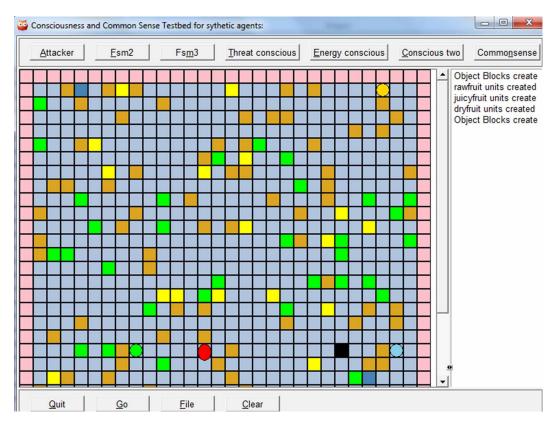
**Figure 1. Animate testbed setup for COCOCA agents**

**Table 1. Energy parameters in animate testbed environment**

| Parameter | Type | Preference | Representation | Affect-value (Units of Energy Given) |
|-----------|------|-----------|----------------|--------------------------------------|
| Rawfruit | Numeric | Energy < 30 | Green Square | 5 units |
| Juicy fruit | Numeric | If encounters on its way | Golden Square | 1 unit |
| Dry fruit | Numeric | Energy > 30 | Yellow Square | 3 units |

capabilities. The percept of an agent is elaborated at each layer by mapping it with different entities to form new associations. This leads an agent to show different levels of intelligence at each layer. The reflexes and reactive behaviors are simple mapping of percept on to behaviors that are implemented through control theory and finite state machines. The deliberative layer and above are the conscious behaviors that are implemented by using Belief Desire and Intention (BDI) models. These layers depict rational behaviors in agents, which have desires based on their beliefs and emotions. These rational desires of agents are treated as implicit or explicit goals that result in control states, which in turn trigger motivations. The conscious actions of agents are the motivated behaviors that satisfy their implicit or explicit goals. Each layer below provides a service to the layer above and hence the functions coded in the lower level layers are invoked and controlled by functions in the layers above. For example, deliberative actions at the deliberative layer use the services of reactive and reflexive layer behaviors to construct deliberated behaviors.

The COCOCA agent's Sense-Plan-Act(SPA) cycle is distributed among six layers and five columns of architecture with different control processes and representations in each layer. The sensed data from the environment is processed in each layer through perception filters for generating different levels of abstractions. The five columns of the architecture are Perception, Affect/Emotion, Cognition, Motivation, and Action.

Information of the external world is perceived in the perception column. If the input stimulus is relating to an alarming event, it overrides motivations of a higher layer and generates an immediate action at the reflexive layer itself. As it is further processed in a higher layer, the association an object forms with other objects is evaluated to see if that object triggers emotions in agents and in turn affects the performance.

In the deliberative layer the precepts are mapped on to agents' Belief-Desire-Intention (BDI) to trigger motivated actions. The attention selector processes in the consciousness layer evaluate the motivated actions in the deliberative layer and update the belief set frequently. The self-reflective layer monitors every conscious action triggered and their effect on the agent's internal state and external world. This forms a feedback for converting some of the beliefs as common sense beliefs. The parameters that may affect the immediate goal or can trigger fear as an emotion in an agent are pushed into global workspace or working memory to get conscious control. The emotions and motivations play a major role in generating consciousness and common sense in an agent's behavior. This leads an agent to manage its motivations and goals by selecting appropriate strategies. The strategies can be either conscious or common sense that is based on the meta-reasoning logic adopted by an agent. The

**Figure 2. Layers of COCOCA architecture**



default meta-reasoning is: if in normal scenarios, common sense triggers and if in fear, consciousness improves. These strategies selected are constructed into action set and sent to an action generator, which in turn changes the external environment (see Figure 3).
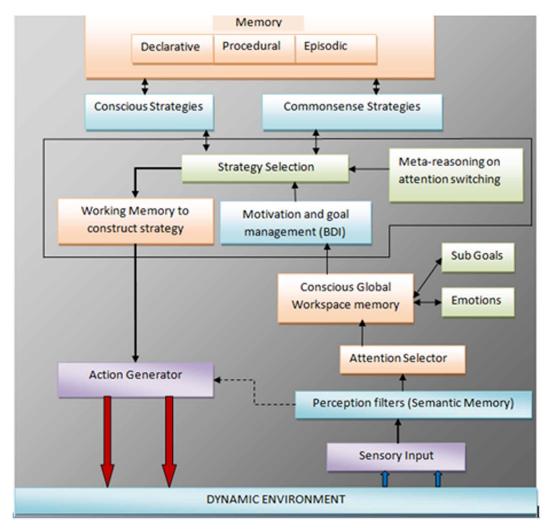
**Figure 3. Cognitive cycle of COCOCA agents**



## Design of Reflexive and Reactive Agents

The reflexive and reactive behaviors of agents are closely associated with sensing subsystem and actuator subsystem. As defined in Sloman architecture (2014), the reflexive layer is defined with a set of ballistic actions that form the response for alarming events in external world. The reflexes are like interrupts that can disable the deliberated actions and execute with higher priority. The reactive behavior in an agent is the awareness about how to react for a particular external/internal event. For example, having the awareness of consuming a fruit to gain energy is a reactive behavior. These reactive behaviors are initiated and controlled by motivation in the deliberative layer, which are triggered to achieve the goals.

The reflexive behaviors are designed by using Finite-State-Machine (FSM), as shown in Figure 4. These are the pre-defined responses for objects in the environment. All agents use FSM as a basic algorithm for movement in the environment. The FSM

defines action for each possible agent state in the environment. The COCOCA agents' internal state can be either static, active, or turned. At any given time, the agent lives in state S, and by executing an action A the agent can move to a new state S' or it may remain in the same state. The algorithm defines the possible action an agent can execute being in a state. It gives the pre-conditions and post-conditions for every action that is being executed. The agent is initially created in a "static" state and changes to an "active" state with an execution of an action. The agent must be in the active state to move in any direction. To take a turn in the corner, the agent must be in the "turned" state. If there are no moves possible in current position, the state of the agent again changes to "static" state irrespective of its current state. If the current state is "turned" and there is free space in next move, its state changes to active again.

## Design of Deliberative Layer

The deliberative behaviors are monitored by control processes in higher layers and are broken down to a sequence of reflexive and reactive behaviors. The processes in this layer use the BDI model for reasoning and planning.
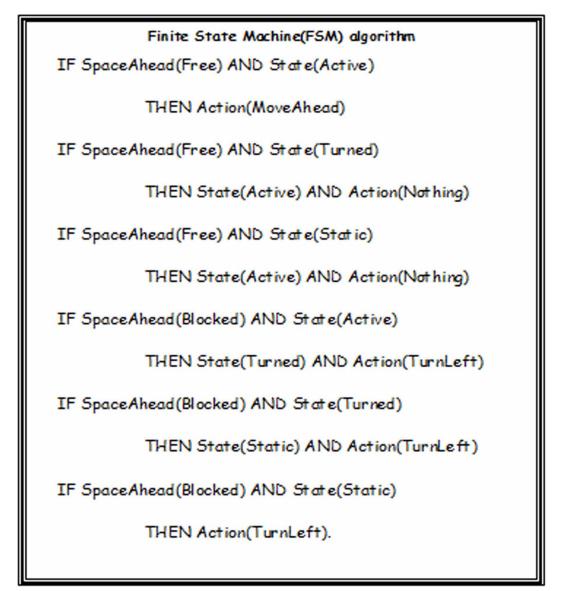
The perceived inputs are processed in this layer to form association with current the set of beliefs. The cross-product of the belief set and desire set gives a set of intentions that are possible in the current state. The conscious process in higher layer weighs the intention set and chooses the one with higher weight as a deliberated action. The belief set of an agent is defined with the facts about the environment and the self. This set initially contains the facts of the external world such as the availability of raw fruits, dry fruits, juicy fruits, obstacles, and border of the arena. The set also includes its internal parameters such as its energy level, state, direction, name, and color. The desire set is a set of actions that an agent can execute based on the actuator set at its disposal. There is a subset of beliefs that are defined as common sense beliefs. Common sense agents may have the same set of desires but use a different set of beliefs for reasoning.

The deliberative cycle of a COCOCA agent shown in Figure 5 explains the BDI model adopted in this layer. This layer maintains an updated set of beliefs, which are continuously cross-verified with parameters in the environment in each cycle. The environment and agent keep changing their states to a new state on each action and in turn their belief sets get updated. The intentional actions are then constructed as a sequence of tasks at lower layers. These intentions in turn belong to either conscious strategy or common sense strategy.
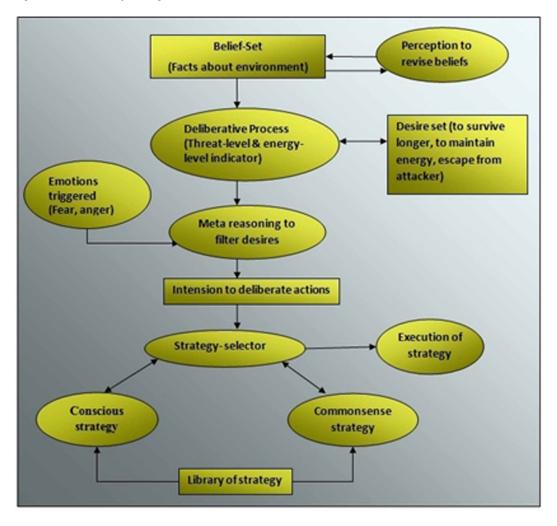
## Design of Conscious Agents

The conscious layer is designed by using the axiomatic (Aleksander, I. & Dunmall, 2003; Aleksander, I., 2007) theory by simulating cognitive abilities that make an agent conscious. The conscious agents are built by using proactive attitudes like beliefs, desires, intentions, and emotions that form the basis for motivated actions, which are conscious by default. The agents are given a collection of strategies that suits different environmental conditions. The conscious agents, based on their goals,

**Figure 4. Finite state machine for reflexive behavior**



Finite State Machine(FSM) algorithm

IF SpaceAhead(Free) AND State(Active)

THEN Action(MoveAhead)

IF SpaceAhead(Free) AND State(Turned)

THEN State(Active) AND Action(Nothing)

IF SpaceAhead(Free) AND State(Static)

THEN State(Active) AND Action(Nothing)

IF SpaceAhead(Blocked) AND State(Active)

THEN State(Turned) AND Action(TurnLeft)

IF SpaceAhead(Blocked) AND State(Turned)

THEN State(Static) AND Action(TurnLeft)

IF SpaceAhead(Blocked) AND State(Static)

THEN Action(TurnLeft).

choose the required parameters, which are in focal attention, from the perception list. The changes in the environment make an agent trigger different actions by choosing different strategies. Each conscious strategy depicts different cognitive abilities to demonstrate consciousness levels in the agents. The conscious agents use internal affect value of objects in the external world and BDI set to choose the strategies. These strategies work on a partial order planning to accomplish the goals assigned. The internal structure of the conscious agent is shown in Figure 6.

**Figure 5. Deliberative layer design**



## Consciousness Layer

The conscious layer cognitive process will have an access to declarative knowledge and procedural knowledge; hence, the agents of this layer will have self-awareness and consciousness of "I know I know" level of Theory of Mind. The control process, on receiving the perceived information from the layer below, will be able to infer possible actions. The mapping is done after a conscious evaluation of each percept and its effect on goal achievement. This layer defines strategies for planning based on the current state of the agent and the agent's preferences with respect to goal achievement. The layer has implicit learning process, a rote-learning for remembering traps after experiencing them.

The agents in the conscious layer are built by using behaviors in the deliberative layer and in-turn in the reactive and reflexive layers. The state-transition diagram for conscious agent is as shown in Figure 7. The agent of this layer can be threat-conscious, energy-conscious, or both. Initially, the agent demonstrates explore behavior
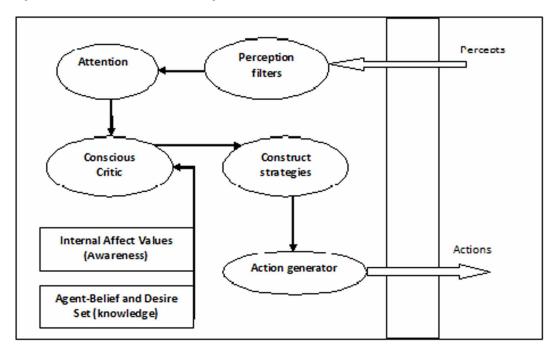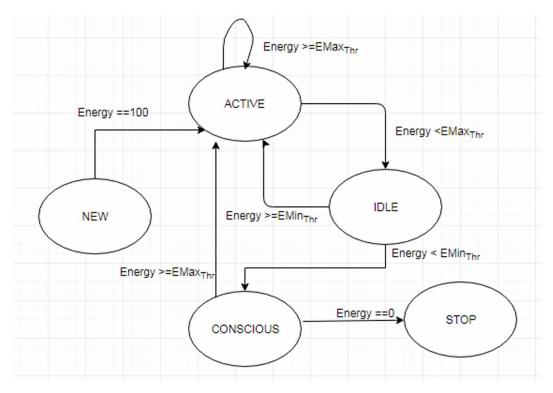
**Figure 6. Internal structure of a conscious agent**



**Figure 7. State-transition diagram for designing conscious agents**

by using simple reflexive behaviors and consume every fruit in the environment. When its internal performance parameters are affected, the agent changes its state to the conscious state. These agents in conscious state evaluate its internal needs and external environment state to choose the optimal strategy. These agents adopt learning methods to repeat the best strategies.

## Designing Energy-Conscious Agents

The initial set of behaviors and actions of agents is:

**$Beh_A$:** { Reactive-Rfruit, Reactive-Jfriut, Reactive-Dfruit, Go-Idle}

**$Action_A$:** { MoveForward, MoveUp, MoveRight,MoveLeft, consume-Rfruit, consume-Jfriut, consume-Dfruit }

**$IP_{iA}$:** { Internal-state, Initial-Energy, Threshold-Energy ($EMax_{Thr}$= 60, $EMin_{Thr}$ = 30)}
where Internal-state can take one of the states in the given set:

$$\{NEW, ACTIVE, IDLE, cONSCIOUS\}$$

$Belif_A$ = { CST1, CST2,CST3,CST4,CST5, CST6,CST7}
Energy consumption in agent: One Move – 1 unit, Idel – 0 units, on trap – 2 units
Agent behavior in the NEW state:
    step1 Create internal state of agent { Energy =100, STATE = NEW}
    step2 Initialize its belief set
    step3 Initialize its action set
    step4 RUN agent
Agent behavior in the ACTIVE state:
    Step 1    Agent with Simple-Reflexive-Explore {Update-Energy}
    Step 2    Adopt default Conscious-Strategy { Update-Energy}
    Step 3    Monitor internal parameters { $EMax_{Thr}$<= Energy <= $EMin_{Thr}$ }
    Step 4    If Energy <= $EMax_{Thr}$ go to IDLE state
    Step 5    Go to step 1
Agent behavior in the CONSCIOUS state:
    Step 1    If Energy <= $EMin_{Thr}$
    Step 2    Initialize strategic-planning
    Step 3    Evaluate Pre-conditions {current energy-level, current position (Grid-location), current precepts}
    Step 4    Change Strategy {Choose next strategy available in list}

```
    Step 5       Monitor Energy
        if   EMax_Thr>= Energy >= EMin_Thr
            i.    Initiate learn-fruit
            ii.   Update knowledge
        Else Go to step 4
Agent behavior in the IDLE state:
    Step 1       Agent in same position for next 5
cycles
    Step 2       If Energy > EMin_Thr AND Energy < EMax_Thr
        Change state to ACTIVE
Else Change state to CONSCIOUS
```

## *Characteristics of Conscious Agents*

- **Attention in Conscious Agents:** The conscious agents have self-awareness of their internal states such as goals, beliefs, and desires. The agents possess access-consciousness and are aware of different food parameters in the environment. For example, if the energy of an agent drops below the minimum threshold, the agent's attention is focused towards fruits in the environment.
- **Attention Switching in Conscious Agents:** The conscious agents are designed to dynamically switch their attention to different parameters to meet the changing needs. The attention switching in an agent can trigger due to a change in either the external event or the internal state.
- **Learning in Conscious Agents:** The conscious agents adopt the rote-learning method to memorize experiences. The agents existing in unknown environments or having incomplete knowledge about the environment face difficulty in reasoning and decision-making in critical conditions.

The learning in agent starts as soon it notices that its internal state is not a desired state and the current strategy is failing to achieve. The agents are given a list of strategies without the knowledge of necessary preconditions to apply them. The agents get a positive reward if the current strategy is affecting the parameter connected with desired state; else it gets a negative reward. The strategy getting a positive reward will be learnt by saving the precondition when it was applied.

A generalized reinforcement learning algorithm given below allows an agent to learn adopt different strategies on different conditions. In learning energy maintenance the set of states, action set and learning parameters are as given below.

```
Set of States S = {Low_Energy, High_Energy, , Current_Energy}
Set of Actions A = { CST1, CST2,CST3,CST4}
(CST1: Consume-only-Rfruit, CST2: Move-opposite-2steps,
CST3: Jump-2steps-ahead
CST4:Consume-only-Dfruit)
Q(s,a) is the cumulative reward gained in previous
strategy adopted (+ve value  if energy increased
```

```
otherwise -ve)
R(s, a') is the reward gained by applying current
strategy
```
$P_{best}$ is the previous best cumulative reward gained for any strategy adopted by an agent

$Best_{Avg}$ is the feasible incremental reward expected after adopting a strategy

```
Step 1: Initialize Q(s,a) =0
     α = 0.1
     Υ = 0.9
```
Step2: Calculate $Q(s',a') + = \alpha * (R(s, a')_{+} \Upsilon * (P_{best} - Q(s,a))$

Step3:    if $Q(s',a') > Best_{Avg}$
```
     Save strategy with precondition
```
$Best_{Avg} = Q(s',a')$

Else if $Q(s',a') < Bes_{Avg}$
```
     Unlearn strategy for current preconditions
     Step 4: s =s' and Q(s,a) =Q(s',a')
     Step 5: go to step 2
```

## Common Sense Layer

The common sense agents are by default conscious agents with respect to state-transitions and behavior preferences. These agents are given some initial common sense strategies, which are similar to beliefs. The agents switch to common sense if they have come across such situations repeatedly in their past experiences. The attackers in the environment create a threat state for agents and survival becomes a decisive goal if there is a threat for life, irrespective of their energylevel; they strive to escape the threat by using their common sense. A similar behavior has been created by giving a common sense belief set that says what plans it can adopt to escape the threat.

The common sense agents adopt known strategies first and then learn to optimize the behavior. The initial strategies are regularly updated by agents through their experiences. The common sense strategies are used by agents when they are aware that they can always escape from the attacker.

### *Common Sense Agent's Behavior in the CONSCIOUS State for Threat Management*

```
Step 1      If  Threat-level = HIGH
Step 2     Initialize strategic-planning
Step 3      Evaluate Pre-conditions
     {Current Threat-level, current position (Grid-
location), current precepts}
Step 4      Choose strategy from common sense by mapping
pre-conditions
```

```
Step 5        Monitor Threat-level
    if Threat - level == MEDIUM
              Continue with the same strategy
         Else if Threat-level == LOW
              Update current strategy as best for
current pre-conditions
         Else if Threat-level == HIGH
              Delete the strategy from list
              Go to step 5
```

## Meta-Reasoning Layer

The meta-reasoning is a layer above the conscious layer of the COCOCA architecture. This layer has self-reflection that monitors and controls the behaviors of agents. The meta-reasoning process monitors the feedback of every conscious strategy of an agent that affects the agent's internal state and external environment. If the desired states of agents are achieved repeatedly by the same strategy in the same pre-conditions, it will be saved to common sense strategies. The emotions due to external events are used by these layers to self-adjust the responses in the next cycles.

## Attacker Strategies

The attacker kills agents for its survival. It adopts different strategies to attack agents when they are alone or in group. The attacker does not kill an agent when the agent is in the safe-zone, which is defined in the environment and when the agents are not in group. The strategy is shown in Figure 8.

## RESULT ANALYSIS OF COCOCA AGENTS IN SIMULATED ANIMATE TESTBED

The results are captured from the agent's behavior in handling their inner states in critical conditions. All agent types are evaluated for energy-level maintenance and escaping rates as performance metrics. Figure 9 is a graph that shows the energy-level maintenance in FSM, energy-conscious, and conscious2 agents in a single-agent environment and in the absence of an attacker. These agents are separately compared for energy levels as they consciously monitor energy levels and use the same belief set. The FSM agent is the least-conscious agent and does not maintain energy levels even when sufficient fruits are available in the environment as it always goes with default strategy.
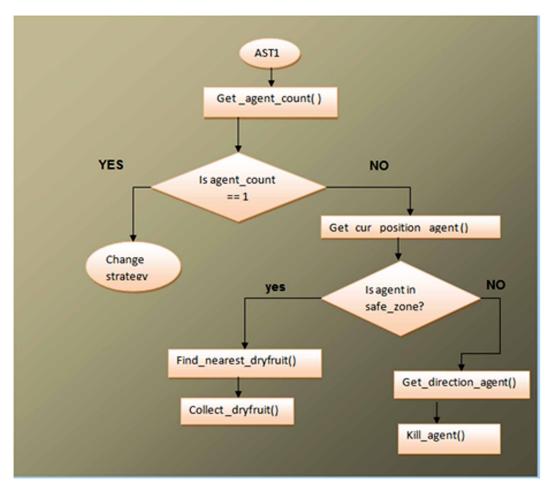
The energy-conscious agent maintains its energy level on threshold, and does not consume raw fruit unless its energy level drops below the threshold. The reinforcement learning in conscious agents enables energy conscious agents to achieve this by dynamically changing the strategies that are more probable to achieve the goal state.

In the presence of an attacker, the agents are evaluated for escape-count. As energy-conscious agents are not threat-conscious, they can be killed by an attacker, but

**Figure 8. Attacker Strategy-1**



during its survival its energy maintenance is better compared to the threat-conscious and FSM agents.

Figure 10 is a graph that shows the number of cycles each agent type survived from an attacker by escaping. The threat-conscious agent could escape for maximum number of cycles as it adopts better strategies. The common sense agents also perform on par with threat-conscious agents by using common sense. The FSM and energy-conscious agents are least performers in the presence of an attacker as they lack conscious awareness of a threat.

The initial distance between the agent and an attacker affects the number of cycles it can survive. If this distance is very small, the agent survives for very few cycles; if not, it can survive a little longer. Table 2 shows the survival cycles and escape count of different agents.
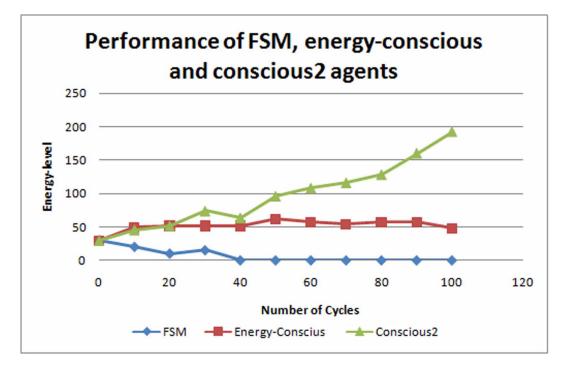
**Figure 9. Comparison of FSM, energy-conscious, and conscious2 agents**
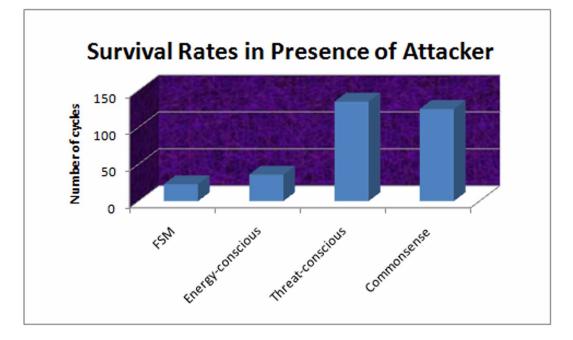


**Figure 10. Survival of agents in the presence of an attacker**

**Table 2. Escape count of agents**

| Agent Type | Number of cycles survived | Number of times escaped |
|---|---|---|
| Threat-conscious | 135 | 60 |
| Conscious2 | 108 | 51 |
| Common sense | 113 | 45 |

## CONCLUSION

The research work highlights the idea that many aspects of consciousness and common sense thinking can be simulated on agents. The research experiment has progressively achieved the results required to justify the outcomes. The conceptual mechanisms of consciousness and common sense have been computationally represented by using cognitive architecture called COCOCA with six-layers-five-columns. Experimentation is conducted by using animate testbed with simulated agents. The agents are made to survive in an environment with attacker coexistence. This makes an agent to consciously focus on the threat level in each move to make a decision for the next move.

The agents of basic layers such as reflexive and reactive are designed by using simple FSM logic. The deliberative agents are built with the BDI model to show motivated actions. The conscious agents tend to maintain desired states according to their native behavior by changing strategies. This process in conscious agents leads to learning optimal behaviors. The meta-reasoning layer upgrades common sense strategies. The agents of common sense layer are the default conscious agents with an initial set of common sense. These agents have pre-hand knowledge for some pre-conditions and adopt these known strategies in critical conditions. The work carried out concludes that a layer of meta-reasoning can build a knowledgebase of common sense through experience.

The comparative study of consciousness and common sense aspects can improve the performance of agents in critical situations. All conscious agents perform better than the FSM agents for both energy-level maintenance and in tackling threats in the environment. The energy-conscious agents are not evaluated for escape from attackers as they lack strategies for escaping. But the threat-conscious agents do maintain energy levels above threshold while simultaneously managing to escape from threats.

# REFERENCES

Aleksander, I. (2007). Machine consciousness. In Velmans & Schneider (Eds.), The Blackwell Companion to Consciousness (pp. 87–98). Oxford: Blackwell. doi:10.1002/9780470751466.ch7

Aleksander, I., & Dunmall, B. (2003). Axioms and tests for the presence of minimal consciousness in agents. *Journal of Consciousness Studies*, *10*(4), 7–18.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1996). ACT: A Simple Theory of Complex Cognition. *The American Psychologist*, *51*(4), 355–365. doi:10.1037/0003-066X.51.4.355

Armstrong, D. M. (1968). *A materialist theory of the mind*. New York: Humanities Press.

Arrabales, R., Ledezma, A., & Sanchis, A. (2009). CERA-CRANIUM: A TestBed for Machine Consciousness Research. In *Proceedings of the International Workshop on Machine Consciousness*, Hong Kong.

Baars, B. J., & Franklin, S. (2009). Consciousness is Computational: The LIDA Model Of Global Workspace Theory. *International Journal of Machine Consciousness*, *1*(01), 23–32.

Baars, F. S., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172. doi:10.1016/S1364-6613(03)00056-1 PMID:12691765

Baars, B.J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baars, B.J. (1998). In the Theatre of Consciousness: Global Workspace Theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, *4*(4), 292-309.

Bargh, A., & Morsella, E. (2008). The Unconscious Mind. *Perspectives on Psychological Science*, *3*(1), 73–79. doi:10.1111/j.1745-6916.2008.00064.x PMID:18584056

Barry Smith, D.W. (1995). Common Sense. Cambridge University Press.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227–247. doi:10.1017/S0140525X00038188

Block, N. (2002). The Harder Problem of Consciousness. *The journal of Philosophy*, *99*. Retrieved from http://www.nyu.edu/gsas/dept/philo/faculty/block/

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, *9*(2), 46–52. doi:10.1016/j.tics.2004.12.006 PMID:15668096

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30*(4). PMID:18366828

Bratman, M. E., Israel, D. J., & Pollack, M. E. (1988). Plans and resource-bounded practical reasoning. *Computational Intelligence*, *4*(3), 349–355. doi:10.1111/j.1467-8640.1988.tb00284.x

Byrne, M.D. (2005). Cognitive architectures in HCI: Present Work and Future Directions. In *Proceedings of the 11ᵗʰ International Conference on Human Computer Interaction*, Las Vegas, NV.

Casella, A., & Giuliani, G. (1999). Scientific thought and common sense. *Presented at the Conference Science as Culture*, Como-Pavia, September 15-19. Retrieved 08/08/2011 from http://fisicavolta.unipv.it/percorsi/philosophy.asp

Davis, D. N. (2001). Control States and Complete Agent Architectures. *Computational Intelligence*, *17*(4), 621–650. doi:10.1111/0824-7935.00167

Davis, D. N. (2002). Computational Architectures for Intelligence and Motivation. In *Proceedings of the 17th IEEE International Symposium on Intelligent Control*, Vancouver, Canada. doi:10.1109/ISIC.2002.1157817

Davis, D. N., & Lewis, S. C. (2004). Affect and Affordance: Architectures without emotions. AAAI 2004 Spring Symposium, Stanford, USA.

Davis, D.N. (2010). Cognitive Architecture for Affect and Motivation. *Cognitive Computation*, *2*(3), 199-216.

Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little, Brown and Company.

Hoffman, D. D. (2014). The origin of time in conscious agents. *Cosmology*, *18*, 494–520.

Minsky, M. (2006). *Emotion Machine*. MIT Press. Retrieved from http://web.media.mit.edu/~minsky/eb4.html

M.V Vijayakumar and Davis D N. (2008). Society of Mind Cognitive Architecture. In *Proceedings of ICAI'08 - The 2008 International Conference on Artificial Intelligence*, Monte Carlo Resort, Las Vegas, NV, July 14-17.

Neumann, V. J. (1946). The Principles of Large-Scale Computing Machines. *Ann. History & Computing*, *3*(3), 263–273. doi:10.1109/MAHC.1981.10025

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Newell, A. (1992). Unified theories of cognition and the role of SOAR. In J.A. Michon & A. Akyurek (Eds.), SOAR: A Cognitive Architecture in Perspective (pp. 25-75). Dordrecht, the Netherlands: Kluwer Academic. doi:10.1007/978-94-011-2426-3_3

Newell, A., & Simon, H. A. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall.

Russell, S. J., & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Shoemaker, S. (1996). On knowing one's own mind. In S. Shoemaker (Ed.), *the first-person perspective and other essays* (pp. 25–49). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511624674.003

Singh, P. (2005). EM-ONE: Architecture for Reflective Common sense Thinking [PhD Thesis]. Retrieved 22 November 2005 from http://web.media.mit.edu/~pus

Sloman, A. (2001). Beyond shallow models of emotion. *Cognitive Processing*, *2*(1), 177–198.

Sloman, A. (2014). The Cognition and Affect Project. Retrieved from http://www.cs.bham.ac.uk/research/projects/cogaff/

Vijayakumar, M. V. (2008). Society of Mind Approach to Cognition and Metacognition in a Cognitive Architecture [PhD Thesis]. University of Hull, UK.

*K.R. Shylaja is currently working as associate professor in the Department of CSE, Dr. Ambedkar Institute of Technology, Bangalore for the last 18 years. She has submitted her thesis on cognitive modeling for consciousness and common sense aspects on artificial agents to Jawaharlal Nehru Technological University, Kakinada, India. Her areas of interest in research are Artificial Intelligence, Cognitive science, Artificial Minds. Published couple of papers in international conferences and Journals.*

*M.V. Vijayakumar is a reformer educationist, and a renowned researcher. He is presently working as a professor (PG, CSE) and in charge for Foreign collaborations in the Dr. Ambedkar Institute of Technology, Bangalore, India. He worked as a Director of Research, Professor and Head of the department of Information Science and Engineering at New Horizon College of Engineering, Bangalore. He completed his BE in Computer Science and engineering from the Department of Computer Science, SJCE, Mysore. He completed an MTech in Computer Science and Engineering from the Department of Computer Science, KREC, Surathkal, presently named as NITK, India. Afterwards, he worked as a software engineer, lecturer, and then he fascinated towards research in his young age. In his interest and curiosity, he joined a PhD Full time course in United Kingdom. He has been awarded PhD in Computer science and engineering, Sub area Artificial Intelligence under the guidance of popular AI scientist Dr. Darryl N. Davis, from the University of Hull, England, United Kingdom. He obtained SIA from London, and Post Graduate Certificate course in Research Training (PGCRT), from the University of Hull, United Kingdom. He worked and has collaborated with many Popular scientists worldwide.*

*E.V. Prasad has thirty-six years of academic experience in technical education and is currently the Director, Lakireddy Bali Reddy College of Engineering, Mylavaram, Krishna District, Andhra Pradesh. Earlier to the current assignment he was the Rector, Jawaharlal Nehru Technological University (JNTU), Kakinada, Registrar and Director of Institute of Science and Technology, JNTU Kakinad, and distinguished himself as Principal and Vice-Principal of University College of Engineering (Autonomous), JNTU, Kakinada. Dr. E.V. Prasad, Director, LBRCE, obtained a PhD in Computer Science Engineering from the University of Roorkee in 1990, an ME in Control Systems from the Madras University in 1978 and a BE in Electronics and Communication from SV University, Tirupati in 1975. He has guided 11 students towards Ph.D degrees, co-authored six books and authored 102 research publications. He visited Malaysia on the invitation of Microsoft in 2004 and Blekinge Institute of Technology, Sweden, as part of MOU between the Universities, in 2010. He received the State Award given to Meritorious Teachers from Government Andhra Pradesh, for the year 2008. He was adjudicated as the 'best teacher' by the students of CSE a number of times. He has proved to be the most successful administrator in making the college campus free of campus unrest by implementing innovative practices and procedures for better academic ambience, aesthetic look, study environment for students and working culture for Teaching & Non-Teaching staff.*

*Darryl N. Davis is a Lecturer in Department of Computer Science, at The University of Hull. He graduated from the University of Sussex with a BSc in Experimental Psychology. He pursued extra-academic interests for the following twelve years. On returning to academia he graduated from Heriot-Watt University with a M.Sc, in Knowledge Base Systems. The MSc thesis was on data mining in a medical domain. He subsequently worked at the University of St. Andrews on human visual perception, particularly human face recognition. He worked in the Departments of Orthodontics and Medical Biophysics at the University of Manchester for his PhD on the use of AI Architectures for Medical Image interpretation, graduating from Victoria University of Manchester with a PhD in Investigative and Diagnostic Medicine. He worked at the University of Birmingham on a number of projects. This includes time as a Research Fellow with Professor Sloman on a project that addressed a number of issues related to the current text. He has been at the University of Hull since 1999, lecturing in Artificial Intelligence and researching in related areas. In 1999 and 2000 he was invited by the Japanese Society for the Promotion of Science to be a visiting professor at the Graduate Institute at Kyoto University. He is a consultant editor (in AI) for Palgrave, MacMillan. His current research interests are diverse with a current focus on architectures for cognition and affect, robotics and agents, machine vision (in particular adaptive segmentation), and data mining. His 120 publications to date include 6 Chapters in Edited Texts, 25 International Journals and 33 International (refereed) Conferences.*