Design of Metacontrol and Metacognition mechanisms in SMCA Using Norms and Affect rules

Venkatamuni Vijaya Kumar, Darryl. N. Davis, and K.R. Shylaja
New Horizon college of Engineering, DR.AIT, Banagalore, India
and University of Hull, UK

Abstract

This research paper investigation concerned with understanding the mechanisms and designing a rules of Metacontrol and Metacognition. This work investigated the concept of metacognition as a powerful catalyst for control, unify and self-reflection. Metacontrol is a part of the Metacognition task. Metacontrol rules decides which deliberative agents are to be learned and ready to perform in different conditions. The deliberative actions are called control action. A Metacontroller determines the relevant control actions. Metacognition is used on BDI models with respect to planning, reasoning, decision making, self reflection, problem solving, learning and the general process of cognition to improve performance. One perspective on how to develop metacognition in a cognitive model is based on the differentiation between metacognitive strategies and metacomponents or metacognitive aids. Metacognitive strategies denote activities such as Metacomphrension (remedial action). Metacognition agents follow well aligned norms, perceptual range, metarules, learning and affect values to maximize its own reward.

1. Introduction

The study of metacognition has grown since the 1970s. In educational psychology, Flavel [6] and others developed a model of children's cognition about a memory (metamemory), understanding (meta-comprehension) and communication (meta-communication). Metacognition is often simply defined as "thinking about thinking" [1, 2, 3]. Broadly defined, metacognition is any knowledge or cognitive process that refers to monitoring and controlling any aspect of cognition. According to Adkins [1] metacognition is thinking about knowing, learning about thinking, control of learning, knowing about knowing. Minsky [7] states that we cannot think about thinking, without thinking about something, where that something is a behavior or activity. The metacognitive act can be referred to as metacontrol. According to Flavel; there are three stages in metacognition: (1) metacognitive knowledge; (2) metacognitive experience; and (3) metacognitive regulation. Metacognitive knowledge contains a database of knowing about an environment, the nature of the task, and strategies used for knowing the facts. Metacognitive experience is after processing a given task, getting knowledge or results. Controlling and (the self reflective) monitoring of progress using cognitive tasks is termed metacognitive regulation [1, 2, 3].

2. Metacognition Architecture

A generic architecture for metacognition consists of three layers (Fig 1): (1) application layer; (2) metacognition; and (3) metacognition architectures. The second layer is called as metacognition consist of three levels: metamemory, metacomphrension and self regulation. Metamemory stores strategies used for executing a task. An execution strategy includes execution parameters: metabolic parameter which includes medicine. The medicine is consumed by the agent to maintain low metabolic rate so that it can survive for longer in the environment. The energy resource parameters are standard processes and knowledge about the environment. Metacomphrension is used for detecting and rectifying the errors.

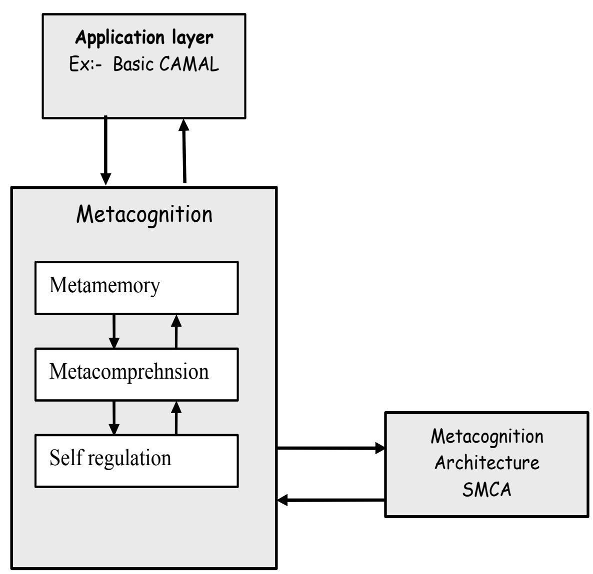


Fig 1. Generic architecture

The self regulation or metamanagement is using for adjusting an error thoughts and to give a feedback.

Adding metacognition concept on top of cognitive architectures, improves the performance. This is similar to updating a systems memory or processor speed of a computer. The architecture remains same but it shows advanced behaviour and functioning [2, 3, 11, 12].

3. Minsky's Meta-model or (Minsky's A,B and C-Brain

Minsky [7] addressed the possible inner mechanisms, and higher level thinking of the mind. Minsky initially postulated an A-Brain and B-Brain mechanisms (Fig. 2).

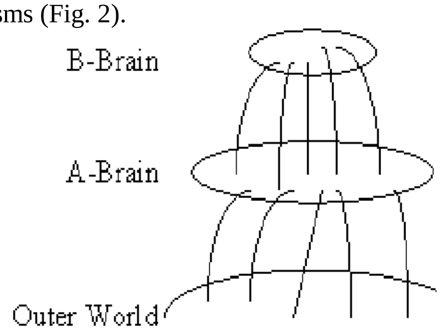


Fig 2. Minsky A B and C-Brain

A-Brain is connected to outer world through sensors and effectors. A-Brian collects information form the outer world or environment. A-Brain will control the cognitive tasks or mental processes in the architecture. The mental processes include perception, memory, imagery, language, problem solving, reasoning and decision making activities. B-Brain act like a supervisor for A-Brian. A-Brain stops or struck or in confusion state to react, then B-Brain makes self reflection of A-Brain. B-Brain can supervise an A-Brain without understanding A-Brain working mechanisms. Minsky suggested that, A & B-Brains can have C-Brain. This can control, watch, and influence the B-Brain. B-Brain and C-Brain works as similar to the A and B-Brains. In addition to this, Minsky suggested "closed loop" concept. The closed loop concept follows transitive mechanism. For example, B is a supervisor of A, C is a supervisor of B and, C is also a supervisor of A. According to Kennedy [6], A and B-Brain's can not mutually monitor and can modify each other. This is called as closed system, but not reflective.

4. Design of SMCA using Metacontrol and Metacognition rules

SMCA [11] is a distributed model of mind, which is an extended architecture of CAMAL [4] with an extra processing layer. SMCA demonstrates planning, reasoning, decision making, self-reflection, problem solving and learning capabilities, predicting, repairing, reviewing, comparing, generalizing and simplifying and many other ways of thinking. SMCA is a six layered architecture with reflexive, reactive, deliberative, learning, metacontrol and metacognition levels (Fig 3). This uses Fungus eater test bed to simulate simple to complex level intelligent agents [11].

Fungus eater Test bed: An agent in fungus eater environment has different biochemical parameters such as metabolism and performance. Metabolism is the rate of energy consumption of an agent to demonstrate simple behaviours. Performance is how optimally an agent can collect the resources in the environment by maintaining its energy levels. The fungus eater environment has various funguses: standard fungus (green squares), bad fungus (black squares) and small fungus (small green square). The goal based parameters are Ore (red stars), Golden-Ore (yellow stars) and Crystals (white stars).This environment is populated with many synthetic artificial agents. These synthetic agents work under the guidance and control mechanisms of higher layers in SMCA [11].

SMCA has many simple agents which work together to exhibit intelligent behaviour. It has many simple reflexive agents to highly intelligent metacognitive agents. The agents at higher level can learn and adapt to the dynamic environments. This architecture simulated six reflexive, seven reactive, fifteen deliberative, nineteen perceptual, fifteen learning, fifteen metacontrol and seventy seven metacognition behaviours.

Reflexive layer: The agents in this lowest level layer can exhibit simple reflexive behaviours. The action of these reflexive agents fully depends on the state of environment given. The reflexive agents can sense the edges of the environment and centre of the environment. These reflexive agents can move in all four directions, if the space is free. It moves to the edges of the environment, if the space is not available or when it is idle. The agents can exhibit two kinds of reflexes (the actions which happen before thinking): simple reflexes which do not require any learning and combined reflexes. The reflexive behaviours are implemented through Finite State Machine (FSM). The FSM is a mathematical model works on finite number of inputs and outputs. This means the agents responds to only a finite set of stimuli. The output of FSM is directly mapped on to the agent's actions.

Reactive Layer: The agents in this layer exhibits a well planned and coordinated actions to satisfy the goal specified for that agent. The reactive agents are controlled by deliberative BDI agents which sets the goal for reactive agents. There are seven reactive agents: Reactive-ore, reactive-golden-ore, reactive-crystal, reactive-resource, reactive-unit, reactive-medicine and reactive-fungus. Reactive-ore agent always moves towards the nearest ore and collect it. Similarly the other agents move towards the intended resources and collect them. Reactive agents understand the affecting parameters of their behaviour such as distance to the resource and the type of resource. As the agent always moves in the direction in which the nearest resource is available. If there are no resources available in their perceptual range they move to the edges of the environment.

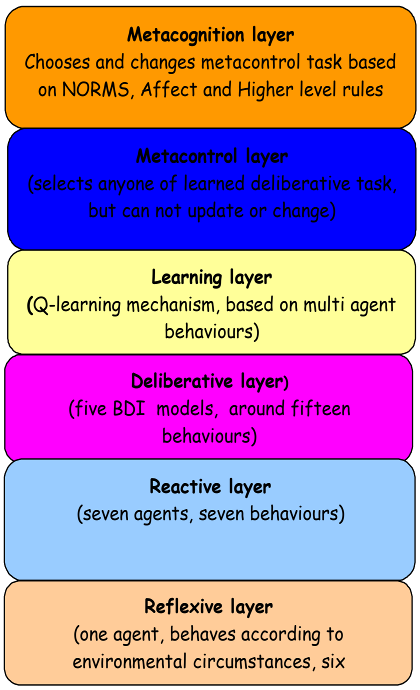


Fig 3. Layers of SMCA

Deliberative layer: The third layer of the architecture is deliberative layer which is populated with many BDI (Belief Desire and Intension) agents. These deliberative BDI agents use required number of reactive agents and reflexive agents to meet the goal specified by the entire architecture. There are five BDI agents: BDI-Ore which selects and controls reactive-ore, reactive-ore-crystal, reactive-fungus and reactive medicine. Similarly it has BDI-Ore-Crystal, BDI-Crystal, BDI-Adaptive, and BDI-Reflexive.

Learning layer: This helps the agents in decision making based on its previous experience or handling the problems at lower level layers. This layer works in control of metacontrol layer. A reinforcement learning algorithm called Q-learning is used for he agent's learning. This algorithm always finds a maximum reward for an action (Fig 4). There are four basic factors contributing in Q-learning algorithm: policy, reward function, value function and model.

Q-learning algorithms work by estimating the values of state-action pairs. The value Q(s, a) is defined as the expected discounted sum of future payoffs. This can be obtained by taking an action 'a' from state's'. Given the delta value from the current state s, selecting an action a, will cause receipt of an immediate goal unit and arrival at the next move. The rules can be symbolic, fuzzy, neural or other, depending on the direction taken in devising the metacontrol and metacognition part of the architecture. The metacontrol mechanisms can be viewed in terms of which the agents use existing controllers, learn behaviours (i.e. existing Q(s, a) values) or learn new behaviours by training the agents.

The **Policy** is a simple function that maps the perception state from the environment to the stimulus-response action of an agent in the given environment. The **reward function** defines the goal for the reinforcement learning, which maps a state and action pair on to a single reward. This determines good or bad events for an agent. The reward function determines the immediate and intrinsic desirability of an agent. The **value functions** define the future payoffs for achieving a reward, being in state's' and taking an action 'a'. It may be sequence of actions taken by an agent to achieve the optimal reward starting from state s over its lifetime. The **model** is used for planning, which helps the agent to decide on an action in a particular state. An agent applies a policy in each move by looking at the value function to maximize the reward. The delta value is calculated from agent's distance and new distance values: Delta is 1 / (Distance + 1); and Delta is Old + ((Distance - New distance)/Distance).

Metacontrol layer: The metacontrol layer is the fifth layer of SMCA. The agents in this layer decide which deliberative agents are to be learned. These metacontrol agents have different levels of skills such as, reflexive, reactive, deliberative and learning capabilities. These learned BDI agents can learn and train themselves to achieve maximum rewards. The metacontrol agents implemented are: learned-BDI-Ore, Learned-BDI-Ore-crystal, Learned–BDI-adaptive, learned-BDI-crystal and learned-BDI-reflexive. Further details of this layer are given in section five.

Metacognition layer: The last layer of the architecture has highly intelligent agents. These metacognition agents' uses metacomponents: norms, affect mechanism and higher level rules to meet the goal set for the entire architecture. It can learn optimal steps taken by an agent using m-norms and affect mechanism by using Q-learning algorithm. This layer details are given in section five.

Q-Learning Algorithm:-

Let Q(s,a) be the expected discount of reinforcement of taking action an in state s, then continuing by choosing actions optimally

1. Initialize a table f with states S, actions A and the Q (utility or reward) value estimates.
2. Select an action a (where a ∈ A) and execute it.
3. Observe the immediate reward r,Reward defined using some agent relation, for example distance to desired object. Observe the new state s', achieved by action a on state s, where a ∈ A and s ∈ S.
4. Update the table entry for Q value using an appropriate rule, for example New(s, a) = Old(s, a) + (r(s) – r(s'))/ r(s).The Q values are nearly converged to their optimal values
5. Update the state: s → s'.
6. Repeat from 2 until learning finished.

Fig 4. Q-Learning algorithm

5. Designing of Metacontrol and Metacognition Rules.

Metacontrol agent decides which deliberative agents are to be learned and ready to perform in different conditions. The deliberative actions are called control actions. A Meta controller determines the relevant control actions. The metacontrol agent learns actions upon the environment. The agent calculates all the combinations of deliberative agent's states (inputs) and actions. Metacontrol agents have different levels of skills, such as reflexive, reactive, deliberative, or learning capabilities.

The metacontrol agent can select and controls any of one of the decision models such as :(1) learned-BDI-ore, (2) learn-BDI-crystal, (3) learned-BDI-ore and crystal, (4) learned-BDI-adaptive and (5) learned-BDI-reflexives. BDI agents should learn themselves by trained method. Hence by adding learning methodology makes more effective. Reward is a goal of the metacontrol agent. This defines the good and bad events for the selected BDI agent. Metacontrol agent's main objective is to maximize the total reward of the running BDI agent. The metacontrol level may be a neural or some Neuro-symbolic hybrid which allows learning. These rules can be used by the metacontrol part of the SMCA architecture. The rules can be symbolic, fuzzy, neural or other depending on the direction taken in devising the metacontrol part of the architecture.

6. Norms and Affect

The affect mechanism defined will make the agent to ensure that its metabolism and energy levels are maintained above the required levels. This makes the agent to check frequently its rate of metabolism and energy levels at regular intervals. In this case it will change its desire to collect medicine. The affect of an agent is calculated by the given rule:

Performance (BDI-Ore) = Ore + Golden-Ore + age

Affect (BDI-Ore) = Norm (BDI-Ore) / Performance (BDI-Ore) Similarly,

Performance (Crystal BDI) = Crystal + Age

Affect (Crystal) = Norm (Crystal)/ Performance (Crystal)

Performance (OreCrystalBDI) = Ore + Golden_ore + Crystal + Age

Affect (OreCrystalBDI) = Norm (OreCrystalBDI) / Performance (OreCrystalBDI).

The main feature of a Metacontrol agent is learning the meaningful behaviours. Q-learning is reinforcement learning method, which is adapted by an agent for learning. Metacomponents affects on the agent behaviour from a sense of what is important instead of what to do. Metacognition agents follow well aligned norms, perceptual range, metarules, and learning and affect values. A well driven agent will maximize its performance as a consequence of learning to maximize its own reward. The metacognition agents can change the framework of BDI agents with reasoning. This level works to control and monitor the deliberative models. Two ways metacognition technique can be applied on humans: controlling and monitoring cognitions; and self reflection of individuals their own mental process. The deliberative models can be switched off or on based on the norms.

The term "norm" is an interdisciplinary term, and can be used to refer to a standard principle or a model used for a right action. The executive processes that controls the other cognitive components are responsible for "figuring out how to do a particular task or set of tasks, and then making sure that the task or set of tasks are done correctly". In a given state an agent before taking an action the agent will compute the norm value. For example if the collected ore is 0, norm of ore is set to 0.75. If it is greater than 0, the norm of ore is computed as total ore perceived divided by total ore collected. Once the norms are computed for each resource, the agent now compares the norm value and adapts the norm with the highest value (Fig. 5).

The metacognition agent checks the affect value of the norm before being adapting. The affect value of each resource is set according to the agents dynamic requirements. If the metabolism rate of an agent is high, the affect value of medicine will be high compare to any other resource. Now the metacognition agent switch the BDI model to deliberative-Medicine and collects the near by medicine. Similarly if the energy level is below the decision-boundary, the affect value of the fungus is set high than any other resource. Now the agent will automatically move towards the nearest fungus, even if the agent's norm is to collect ore or any other resource. This ensures that the agent survives for longer in the environment.

Norm 1

Collected (Ore) = 0,

Norm (OreBDI) = 0.75.

Collected (Ore) > 0,

Norm (OreBDI) = Perceived (Ore)/Collected (Ore).

Collected ore is Ore1 + Golden ore1, Collected (Crystal) = 0,

⇒ Norm (Crystal) = 0.75.

Collected (Crystal) > 0,

Norm (Crystal) = Perceived (Crystal)/Collected (Crystal).

(Collected (Ore) + Collected (Crystal) = 0,

Norm (Ore_CrystalBDI) = 1.

(Collected (Ore) + Collected (Crystal)) > 0,

Norm (OreCystalBDI) = Perceived (Cystal+Ore)/Collected (Cystal+Ore).

Fig 5. Norm1 used in Metacognition agent

The metacognitive agent first uses its perceptual range to increase its belief set and draws an intension. It uses affect mechanisms to decide on need of metabolism or need of food. Then it selects the norm to decide which BDI model to use. If this action is not leading to optimal results, it takes remedial actions by switching between the BDI models. It can also learn optimal steps taken by an agent using m-norms and affect mechanism by using Q-learning algorithm.

The affect mechanisms are used by norms to improve the agent's performance by switching between BDI models as and when required, depending on the agent's dynamic requirements. The algorithm given below shows how a metacognitive agent works.

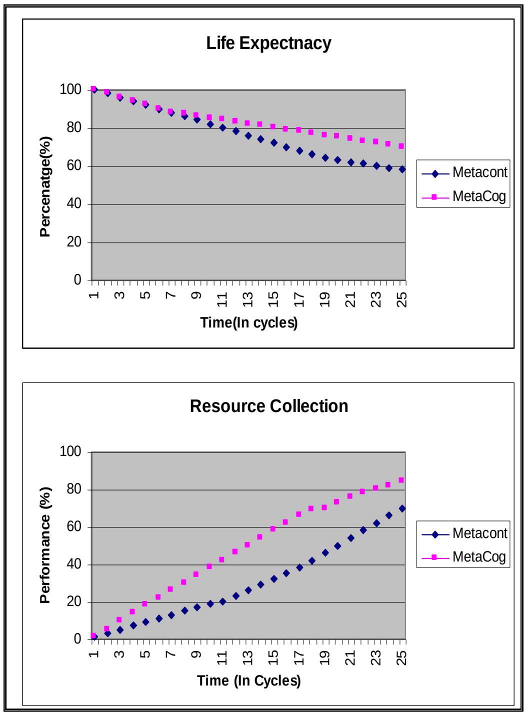
Algorithm to implement Metacognition agent:

1. Map Internals states onto Belief Set from the perceptual range or perceptual level increases the agent's belief set for sensing in an environment. Example level 5 returns Sense List = [5-spacefree, 4-agent, 3-spacefree, 2-fungus, and 1-spacefree].Updates Belief Set with Perceptions and perceptual range.
2. Use Affect mechanism (metacomponent), to find a need of the Metabolism and need of a food.
3. Use a Metacomponets such as Norms or M-Norms (Such as Norm1, Norm2, Norm3, ETC are standard rules) to decide which BDI model to choose in right time by using right decision (optimal decision)by comparing resources available and balance the resources in a testbed.
4. Metacomphrension or Remedial action: Select Appropriate Belief-Desire-Intention Combination (BDI-Ore, BDI-Crystal, BDI-Ore Crystal, ETC), by comparing the architectural results.
5. Metamanagement: Uses M-Norms to Switch the BDI Models (Such as BDI-Ore, BDI-Crystal, BDI-Ore Crystal, ETC),
6. Schema training: Use a Q-Learning (given below) for optimal steps taken from agent by using M-Norms and Affect Mechanism (mtacognition level).
7. Repeats the steps (Step1 to Step6) until Simulation ends.

7. Simulation Results

Experiments were conducted separately for each type of agent. In order to compare results in the experiment, the same statistics were collected.

Different types of agents in different level of architecture were employed for these experiments. The metacontrol and metacognition agents began the experiment with the same percentage (100%) of life expectancy, and resources (refer Graph 1). The metacognition agent manages an energy level of 72%, compared to 58% from the metacontrol agent. The metacognition agent collected 82% of resources, and the metacontrol agent collected 68% of resources. The metacognition mechanism agent performs better than the metacontrol agents. This explains through result graphs, how the concept of metacognition improves the performance through unification.



Graph1. Metacontrol v/s Metacognition Agent

8. Conclusion

This research paper illustrated how to design and adopt a metacontrol and Metacognition mechanisms for cognitive architectures in the broad area of Artificial Intelligence. This paper also given how to frame rules for Metacontrol and Metacognition Mechanisms using Norms and Affect models.

The result concludes that BDI with Metacognition agents are better than other agents. A Metacognition agent collects more resource and manages the higher life expectancy than all other agents.This result proves a concept of metacognition is a powerful catalyst for control, unification and self-reflection. Metacognition used on BDI models with respect to planning, reasoning, decision making, self reflection, problem solving, learning and the general process of cognition improved the performance. This research paper explained how mind research can benefit from the use of metacontrol and metacognition. Finally, this research gives a clear roadmap for researches, to develop a metacognition and metacognition concepts on cognitive modeling.

9. References

[1] J.Adkins Metacognition, university of Saskatchewan, Canada.

[2] M.T Cox, Metacognition, problem solving and aging, Cognitive Science Tech. Rep. *Atlanta, Georgia Institute of Technology*, 1994, College of Computing No. 15.

[3] M.T. Cox, *Metacognition* in Computation: A selected research review, *Artificial Intelligence*. *169*(2), 2005, 104-141.

[4] Davis D N, 2007, *Linking Perception and Action through Perception, Experimental & Theoretical Artificial Intelligence*, and JETAI is available online at: http://www.tandf.co.uk/journals/tf/0952813X.html.

[5] Flavell J H, 1979, *Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry*. American Psychologist, *34*, 906-911.

[6] Kennedy C M**,** 2003, “Distributed Reflective Architectures for Anomaly Detection and Autonomous Recovery", PhD thesis, University of Birmingham.

[7] M. Minsky, The *society of Mind,* (Cofounder of the artificial intelligence, Laboratory, MIT, 1985).

[8] M.Minsky, A Response to Four Reviews of the Society of Mind," Artificial Intelligence, Vol. 48, pp. 371-396, April 1991.

[9] Raja A**,** 2003, Meta-Level Control in Multi-Agent Systems, PhD Thesis, University of Massachusetts Amherst, http://www.sis.uncc.edu/~anraja/thesis.pdf, September 2003, [Accessed 15th May 2005**].**

[10] Stillings N, Cognitive Science. Second edition. Cambridge, MA: MIT Press, 1995.

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

[12] M.V Vijayakumar and Davis D.N, July 2008, Principles of Artificial Economics in Animal Minds, AIPR-08, Orlando, FL, USA.

[13] M.V.Vijayakumar and Davis D N, 2008, Society of Mind Cognitive Architecture, Accepted ICAI'08 - The 2008 International Conference on Artificial Intelligence, Monte Carlo Resort, Las Vegas, Nevada, USA (July 14-17, 2008).

[14] Zalta E N, 2005, *Stanford Encyclopaedia of Philosophy [online]*, available: www.platostanford.edu,