SiMA-C: A Foundational Mental Architecture

Samer Schaat
Institute of Computer Technology, TU Wien, Vienna, Austria
samer.schaat@tuwien.ac.at

Abstract
Cognitive architectures as frameworks for integrated computational models of the mind often focus on human reasoning capabilities, and sometimes are extended by theories of emotion. The SiMA-C approach starts with low-level mental processing and bases high-level process on its low-level foundation. With the aim of developing a unified model with functional equivalence to the human mind, descriptive concepts of perception, motivation, emotion, and cognition are translated into a functional model of data activation, valuation, mediation, and evaluation. The resulting SiMA-C mental architecture solves the problem of adapting an agent’s decision on the current internal state (demands from bodily needs and norms) and the external environment (affordances) and hence mediates between them. Using the SiMA-C model in simulations of environmental-friendly consumer decisions gives an example of a concrete application.

Keywords: Mental architecture, decision making, simulation

1 Introduction
With minds as information processors, one approach of representation is given by computational models. The objective of such approaches can be to get insights of the human mind (Cognitive Science) or to build intelligent technical systems (Artificial Intelligence). We think that the latter is best approachable by the former. Therefore cognitive architectures are developed, which can be defined as “a domain-generic computational cognitive model that captures essential structures and processes of the individual mind for the purpose of a broad (multiple domain) analysis of cognition and behaviour” [14, p. 33]. The currently most prominent approach to reach this objective is to use symbolic systems (cognitivist approach), in particular production systems (e.g. ACT-R [1]), which focus on symbolic knowledge processing that can be represented as if-then rules. Such systems often focus on high-level cognition, aiming to explore human capabilities that supposedly separate them from other animals [2, pp. 187ff.], and hence neglect the (evolutionary) foundations of human cognition. However, various psychological and neuroscientific findings (e.g., [6], [4]) emphasize the key role of unconscious mechanisms, in particular emotions and desires for human decision making. Seminal research even point out that reasoning (including decision making) is based on emotion (e.g., [8][9]), and provide a hierarchical concept that connect the different levels of life-regulation (metabolic...
regulation, reflexes, drives, emotion, feelings) [6]. Some cognitive architectures consider the challenge of using motivations and emotions as building blocks, such as [3] and [13]. However, an holistic computational approach that integrates and unifies these different levels is still an open challenge in Cognitive Science and AI. Such a model can be regarded as a foundational mental architecture.

The SiMA (Simulation of Mental Apparatus and Applications, formerly known as Artificial Recognition System) research program offers a point of departure for considering the mentioned challenges in cognitive architectures. The SiMA approach was initiated by Dietmar Dietrich [7] due to recognizing the lack in AI research to solve problems that humans cope with intuitively on a daily basis, like recognizing dangerous situations. With the premise that the human control system is currently the best system able to cope with complex problems, the SiMA approach aims to model and integrate unconscious and conscious aspects of the human mind, instead of solely focusing on reasoning capabilities as often done in AI. The developed SiMA model is used to explore basic human capabilities in Artificial-Life simulations [10] [11]. Analyzing the functionality of the SiMA model enabled reducing it to a lightweight model with higher degree of unification, called SiMA-C (SiMAC-compressed), which is described here in detail for the first time.

Identifying possibilities for unification and simplification was not only a result of using the SiMA model as a test-bed in Artificial-Life-Simulations and observing the impact of model variables, but is also the result of applying the SiMA model in recent decision-making simulations. Of course, the requirements of developing a model for basic research (SiMA model) is also different than the requirements for an application-oriented model (SiMA-C).

2 Model Overview: Adaptive Decision Making

The basis for unifying the functionality of an artificial mind is to focus on its main objective: to mediate between the internal and external world, to fulfill the different (contradicting) conditions posed by the internal and external world. Information about the agent’s current state, which is able to integrate the different demands and affordances from the internal and external world, serves as the core of an unified model.

A central question a model of decision making has to answer in this regard is how to choose a goal that improves the agent’s state best, in short and long-term. The key assumptions of the presented model is that activation and valuation are sufficient foundational functions to choose relevant goals for the adaptation on demands and affordances from the internal and external environment, and to mediate between them.

The SiMA-C model solves the problem of goal selection based on valuated memories that are activated by conflicting internal and external sources, representing demands and affordances, and possible reflections of this selection using emotion as a representation of an agents current state of pleasure, displeasure, and conflict. Model processing is triggered by a change in bodily needs or the external world, which causes data activation by two sources: Demands and affordances (see Figure 1). Hence the activation process determines possibly relevant data for the current internal or external situation.

Two types of demand sources are distinguished: physical (bodily demands represented by drives and pain from body perception) and psychological (activated memorized norms). Both demand sources activate memorized goals and norms - directly by triggered activation or indirectly by spreading of activation - that are expected to bring pleasure by satisfying these demands. Goals may also be memorized with expected displeasure, if an (activated) goal would prevent the fulfillment of a demand or if a goals object is expected to bring harm.
Beside demands, affordances are the second source of data activation. They activate memories that are similar to the current external world and - indirectly via spreading of activation - norms that are valid for the current external situation.

![Figure 1: SiMA-C model overview.](image)

The relevance of activated goals for the current internal and external situation is determined by different valuation processes, triggered by the different activation sources, and considering the goal’s memorized valuation for these sources. Hence a goal may have different valuations of expected pleasure and displeasure regarding the different activation sources. The current relevance of the single valuation is determined by the caused activation from that source. Besides memorized normative valuation, activated by normative sources, and bodily valuation, activated by bodily sources, perceptive valuation, activated by external perception is distinguished. The first two types of valuations correspond to an expected short-term fulfillment of bodily or normative demands, the latter uses a goal’s memorized summary valuation, which provides integrative holistic information about all aspects of how a goal changed the agent’s state, considering long-term expectations and context. The resulting data from the valuation process can be termed effective or affective, in an etymological sense, meaning that this data should have an impact (on the decision).

Contradicting valuations in a goal (e.g., expected pleasure for normative demands, but expected displeasure for bodily demands) cause conflicts. Additional conflicts are caused by contradictions between demands and affordances or be-tween contradicting goals. For example,
if an internalized norm is activated, goals unrelated to the norm would prevent its fulfillment. Therefore, they would be marked by a conflict. Two types of conflicts are distinguished: normative and reality conflicts, with ownership conflicts being a specialization of the latter. The different kind of conflicts are addressed in processes of mediation, which operate by changing the different valuations in a goal. The result can be called arranged data.

Parallel to these processes, the generated displeasure from the different demands and activated data, the currently experienced pleasure and conflict intensity together form the agent’s state indicator, representing the agents current emotional state. A share of the displeasure from drives is used as so-called neutralized intensity for reflective processes, regulating the grade of dual processing.

Based on the reactive processes described so far, evaluative reasoning processes, which relate the different valuations to the state indicator, are possible. The separation between valuation and evaluation corresponds to a dual processing model of the human mind. The degree of evaluation is dependent on its necessity in case of ambiguities between goals, and the agent’s neutralized intensity. The overall process corresponds to weighing the different valuations in integrating them to a single relevance value. This evaluation corresponds to a multi-criteria aggregation aiming for displeasure minimization and pleasure maximization, while considering a goals conflicts and the agents confidence in the valuations. However, the guiding principle is that of a satisficing, not an optimizing agent (cf. bounded rationality [12]). Overall, the evaluation process results in determining the most relevant goal for the current internal and external conditions.

Overall, the adaptive functionality of the control unit can be generated by valuation, mediation, and evaluation of activated goals. These functions are further described and exemplified in the next sections.

3 Activation

All memories are activatable. Conceptually an activation can be regarded as a requirement of adaptation: Different activation sources that require the agent to adapt on new internal and external situations activate appropriate memories. The stimuli of activation sources are percepts, drives, and emotion. They cause a trigger activation process, which matches activation sources of the external and internal situation (drives and emotions) with memories. Triggered activation, then, is spread through associated memories. In both cases, triggered or spread activation, the current activation value of data is determined by the previous activation value, the current activation intensity, and activation source’s salience.

In case of triggered activation, the activation intensity $a$ is given by the matching value (between activation source and memorized data), weighted by the source’s salience. In case of drives as activation source, the salience is given by the quota of affect; for emotions the salience is given by their intensity (derived from their displeasure and pleasure values); the salience of percepts of the external world are given by the agent’s focus of attention, controlled by various factors, e.g., by psychophysics (e.g. loud colors) or the stimuli sources.

Drives directly activate goals that are known to fulfill the underlying bodily need. Information about the bodily source (drive organ and drive component) is used as stimuli features of drives. The activation is determined by memorized ability to reduce the need.

External Percepts may activate memorized objects or actions (and hence, indirectly, norms). Stimuli features of an object are multi-modal attributes, and the matching value is determined by the number of similar qualitative attributes (e.g. color and shape). An action’s identifying features are the action subject and action name.
Emotion activates memorized objects. Stimuli features are emotion factors of the memorized emotion associated with an object. Matching of these determines the activation value.

Hence goals and norms are only activated indirectly by spreading of activations. After memories are trigger activated, their activation value spreads to associated memories. Models of spreading activation represent an approach to memory retrieval in a generic network architecture [7]. In the SiMA-C model the activation value set by trigger activation is used as an initial output value for all associated data. The input value, then, is dependent on the received activation value, the previously received activation, and the association strength.

In case of spread activation of norms, activation is constrained to descriptive norms (identified by their type, see Figure ??). Hence two kinds of norms are distinguished based on the conceptual separation of injunctive and descriptive norms (e.g. [5]). However, the separation is only implicitly considered in two different activation possibilities of norms. Injunctive norms are activated by the current situation through triggered or spread activation - as any other goal. Descriptive norms are activated if their execution is perceived in the external world. In particular, activation happens in the process of salience activation.

Overall, due to the activation process all possibly relevant data for the current situation are processable. Every activated goal is subsequently considered as a possible and appropriate goal for the current situation. Overall, the activation process impacts decision making only via the next process that it triggers: goal valuation.

4 Valuation

The essence of memories is to form expectations about the purpose and effects of goals, i.e., objects and actions, providing information about their value for the agent. These expectations are updated with the goal’s feedback on the agent’s bodily and normative demands.

The essence of valuation is considering these memorized feedback and adapt the resulting expectations on the demand’s current relevance and the goal’s applicability. Every goal has memorized valuations representing information about the expected purpose for the agent. Hence a memorized valuation is the assignment of an displeasure and a pleasure value to a goal that (1) harmed the body or satisfied a bodily demand (bodily valuation), (2) fulfilled a commandment or broke a prohibition (normative valuation).

These valuations are short-termed, i.e., they value the direct feedback associated with actions and objects. In addition to that, every object has a memorized summary valuation. It provides an integrative holistic information about all aspects of how an object changed the agent’s state, directly or indirectly. The summary valuation subsume all (goal) experience the agent had with it and considers “the whole picture” in providing a summary valuation: it integrates the impact on bodily and normative demands, and their conflicts; long-term feedback, the context. Besides providing information about an object’s purpose, it can be seen as an indication for general approach or avoidance.

Valuation is the process of determining the expected relevance of goals for the agent’s current bodily, normative, and perceived situation. If possible goals are activated, a new valuation process is triggered that weight the memorized valuation with the current activation value to determine the current relevance of the goal. That is, only goals are valuated, since valuation is for doing. However, since goals consist of and are activated - amongst others - by associated objects, their relevance is also considered.

Valuation is done depending on the corresponding activation source type (percept, norm, drive, emotion). Goals have memorized normative and bodily valuations and are associated to objects, which have memorized summary valuations. If a goal is activated, memorized
valuations are weighted with the goal’s activation value from the corresponding activation source type. In case of activation from drives, the goal’s memorized bodily valuation is weighted with the corresponding activation values. If a percept or emotion activates an memorized object, the associated summary valuation is weighted. In case of activation from norms, the goal’s memorized normative valuation is weighted. But since norms are themself only activatable by other valuation sources a normative valuation always implies at least one other type of valuation (bodily or summary valuation, see Figure).

Hence the impact of memorized valuations is dependent on the activation value, which represents the relevance of data for the current situation. Therefore in the process of spreading of activation information about the valuation sources is considered.

![Figure 2: Activation and valuation in SiMA-C.](image_url)

After valuation is conducted, a goal’s *valuated relevance* is set (or extended in case of previous valuations). Following the principles of the primary process, displeasure and pleasure (i.e. the pros and cons) are not weighted against each other, but only the dominant one is chosen to set the relevance value. No-dominance of displeasure or pleasure would increase the goal’s ambiguity, which increases the necessity for evaluating the goal, where both factors (amongst others) are considered (see Section 6).

## 5 Conflict Mediation

(Simulated) humans are conflicted beings, since the human mind is not able to satisfy the various conditions posed by different (and often contradicting) demands and affordances. Conflicts may occur between different goals, but also within a goal in case of contradicting valuations (e.g., dominating displeasure in bodily valuation of a goal, but pleasure in its normative valuation). All conflicts occur dynamically in the process of the various inter-playing decision factors. Hence opposed to valuations, conflict values are not memorized (expect in summary valuation).

The main source of conflicts is normative. All activated norms demand their fulfillment. If a norm is activated and not fulfilled by a goal, i.e., no association between the norm and the goal exist, the goal is marked with a norm conflict.
Another source of conflict is a discrepancy between wish and reality. If a goal is physically executable (i.e., not a strategic goal) and highly valued due to current bodily and/or normative demands, but the affordances of the current situation do not enable its execution, the goal is marked with a reality conflict. Due to conflict mediation (see below) and possibly already in spreading of activation, associated goals that are executable in the current situation are activated.

Conflict occurrence triggers mediation. An example for a widely used mediation mechanism is sublimation. If an agent has learned that two goals are able to satisfy demands similarly, they are associated in memory. Additionally conflicts are considered as weighting factors in the evaluation process (see Section 6). Different mediation mechanisms are possible, dependent on the agent’s configuration (i.e., personality), conflict type and intensity, and the agent’s current ego-strength (represented by the neutralized intensity).

6 Evaluation

The SiMA-C model is a dual processing model of the human mind. The two processes can be regarded as unconscious and conscious or as primary or secondary process. However the second is dependent on the primary process and is only a extension of it. Both follow different principles, but operate on the same values. The primary process activates and valuates data and tries to mediate in case of conflicts. However these processes only consider local conditions, e.g., valuations according to a specific demand. If necessary and possible, the second process uses a global view in considering all factors. After the primary process’ memory-based operating principle, the secondary process uses reflective operations in explicitly relating the agent’s current state to the expected state represented by the goal’s valuation.

Overall goal selection in the SiMA-C agent is based on a goal’s relevance value. The determination of this value is a stepwise process. In the primary process valuation and its manipulation sets the relevance value, which may be extended in the secondary process by evaluation. The used terminology reflect the functionality of the processes of valuation and evaluation, and also emphasizes their relation: evaluation includes and extends valuation. This corresponds to the assumptions that the secondary process can not change the primary process’ goal’s valuation, but only uses it differently (other operating principles). This is immanent to the concept of the unconscious in the primary process.

According to the principles of the primary process the different valuations are just aggregated to a valued relevance. In the secondary process evaluated relevance is calculated in weighted multi-criteria aggregation. Evaluation is able to use all information gained so far in the primary process to decide the most relevant goal in the current situation, in particular to decide which goal is overall best for enhancing an agent’s state. However, in considering bounded rationality, evaluation follows a approach of satisficer, not optimizer. In this regard evaluation is done gradually depending on two dynamic factors (see Figure): (1) the necessity for evaluation represented by the clarity of goals’ relevance (i.e., small distance between goals with the highest relevance), and (2) the possibility represented by the agent’s neutralized intensity.

7 Conclusion

The SiMA-C model shows how a unified model of data activation, valuation, mediation, and evaluation is able to provide a functional representation of motivation, emotion, and cognition in integrating them. In particular it provides mechanisms to adapt decisions on the agent’s
possibly contradicting current internal and external state. The mind’s dualistic character is represented by distinguishing between valuation and evaluation, which use the same variables (activation intensity, pleasure, and displeasure) in different manner. The application of the SiMA-C model in a social media scenario demonstrates its ability to explain environmental-friendly decisions.

References


