

# *Text comprehension and the computational mind-agencies*

**Romi Banerjee & Sankar K. Pal**

**Natural Computing**  
An International Journal

ISSN 1567-7818

Nat Comput  
DOI 10.1007/s11047-014-9478-x



**Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Text comprehension and the computational mind-agencies

Romi Banerjee · Sankar K. Pal

© Springer Science+Business Media Dordrecht 2015

**Abstract** Guided by a polymath approach—encompassing neuroscience, philosophy, psychology and computer science, this article describes a novel ‘cognitive’ computational mind framework for text comprehension in terms of Minsky’s ‘Society of Mind’ and ‘Emotion Machine’ theories. Observing a top-down design method, we enumerate here the macrocosmic elements of the model—the ‘agencies’ and memory constructs, followed by an elucidation on the working principles and synthesis concerns. Besides corroboration of results of a dry-run test by thoughts generated by random human subjects; the completeness of the conceptualized framework has been validated as a consequence of its total representation of ‘text understanding’ functions of the human brain, types of human memory and emulation of the layers of the mind. A brief conceptual comparison, between the architecture and existing ‘conscious’ agents, has been included as well. The framework, though observed here in its capacity as a text comprehender, is capable of understanding in general. A cognitive model of text comprehension, besides contributing to the ‘thinking machines’ research enterprise, is envisioned to be strategic in the design of intelligent plagiarism checkers, literature genre-cataloguers, differential diagnosis systems, and educational aids for children with reading disorders. Turing’s landmark 1950 article on computational intelligence is the principal motivator behind our research initiative.

**Keywords** Society of mind · Thinking machines · Reflective cognitive architecture · Concept-granulation · Natural computation · Artificial general intelligence

## 1 Introduction

Reading furnishes the mind only with materials of knowledge; it is thinking that makes what we read ours.—John Locke

The world isn’t just the way it is. It is how we understand it, no? And in understanding something, we bring something to it, no? —Yann Martel, *Life of Pi*.

‘What is the mind? What is thinking? How does the mind granulate, associate and summarize concepts? How does the infant-mind ‘understand’ and develop language-skills? Is there a generic procedure underlying the functioning of the mind? If yes, can we define it in computational terms?...’—enigmas that always have and yet continue to baffle philosophers and scientists, alike.

Alongside philosophical discourses on the origin of the mind, recent developments in cognitive science—integrating experimental and theoretical investigations across neuroscience, psychology, linguistics and artificial intelligence—and technologies that help probe into the inner brain-activities, present today the practical complexities in pursuing investigations on the above questions. Surprisingly, the intricacies are yet to deter researchers from probing into the working of the mind.

It was while we were attempting an integration of the computing with words (CWW) (Zadeh 1996), natural language processing (NLP) and affective computing paradigms (Picard 1997) towards a methodology of text

---

R. Banerjee (✉) · S. K. Pal  
Center for Soft Computing Research, Indian Statistical Institute,  
Kolkata, India  
e-mail: rm.banerjee@gmail.com

S. K. Pal  
e-mail: sankar@isical.ac.in

comprehension in Banerjee and Pal (2013), Pal et al. (2013) that questions on how does the human mind recall, visualize, granulate and associate perceptions—despite information insufficiency or ambiguity—to form a universe of thoughts (Pinker 2007); identify affective, rhetoric and prosodic elements in text; measure comprehension, etc., intrigued us and prompted the formulation of the concepts illustrated herein.

This article describes our efforts at defining a framework of a cognitive computational mind—an abstraction of the human mind, formed by assimilating different, dynamic and co-operative intelligent components, as do components of the brain or body, giving rise to appropriate emergent structures and dynamics. Our focus here lies exclusively on a computational mind as a text understander.

Referring to the parts of the brain and their functions towards language comprehension (Price 2000; Ramachandran and Blakeslee 1999), we endeavor enumerating a ‘society (Minsky 1986)’ of self-evolving and self-organizing modules (or ‘mind-agencies’) to form a system capable of mimicking each of these brain-functions. A system where the sum of the complex individual functions of the modules would result in a granule of comprehension, quite indistinguishable from the thought-components that lead to it—embodying the basic philosophy of the granular computing paradigm (Lin 1997; Zadeh 1998).

Granular computing is the manifestation of the human ability to perceive the real world across multiple levels of abstraction or granularity—the process of extraction, grouping and manipulation of concepts into hierarchies of coherent modules that fit a given context. It is by processing these different levels of granularity that the mind arrives at associations between interdisciplinary knowledge elements, leading to a greater understanding of the world. Granular computing is thus, an innate human problem solving mechanism and consequently a significant intelligent system design tool. The philosophy of granular computing is rooted in the principles of grouping (Todorovic 2008; Wertheimer 1923) of Gestalt Psychology (Kofka 1935; Todorovic 2008; Wertheimer 1923)—motivating rules of organization of micro-perceived scenes into a complex visualization—a ‘Society of Mind’ approach to the construction of granules of perception where the ‘whole is other than the sum of the parts (Kofka 1935)’.

... In my theory the analysis is based on many interactions between sensations and a huge network of learned symbolic information. While ultimately those interactions must themselves be based also on a reasonable set of powerful principles, the performance theory is separate from the theory of how the system might originate and develop... Thinking always begins with suggestive but imperfect plans and

images; these are progressively replaced by better—but usually still imperfect—*ideas*.—(Minsky 1975).

Post, a brief literature survey of the popular existing models of the human mind (Langley et al. 2009; Singh 2003), we chose Minsky’s ‘Society of Mind (Minsky 1986)’ and ‘Emotion Machine (Minsky 2006)’ theories as the foundation pillars of our work, for a number of reasons. These theories—

- (a) Are implicitly built around Gestalt’s Psychology principles, and in turn the concept of granulation, as is evident in the undertones of the quote above, and in their acknowledgement of Max Wertheimer’s concepts in Wertheimer (1923) of ‘productive’ (intuitive, commonsense-based) and ‘reproductive’ (learned, deliberative, reflective, self-reflective, self-conscious) thinking.
- (b) Covers the entire spectrum of views on the philosophy of the mind, from the ‘dogma of the Ghost in the Machine’ of the intellectualist legends as well as the more practical views of Ryle’s in (1949),
- (c) Inherently recognizes the ‘fast and slow thinking’ (Kahneman 2011) processes, and
- (d) We were particularly challenged by the fact that since its inception, while the ‘Society of Mind’ has been widely used (Baars 1988; Franklin 2003; Kokinov 1989, 1994; Majumdar and Sowa 2008; McCauley et al. 2000; Zhang 1998), the ‘Emotion Machine’ has seen sparse implementation initiatives (Morgan 2010; Morgan 2013; Singh 2005).

Besides Minsky’s ideas, our work draws key inspirations from natural language understanders designed over the last four decades. Turing’s landmark paper (Turing 1950), the year 2012 being named the ‘Alan Turing Year’ and that we are yet to design a machine that wins the ‘imitation game’, despite it being six decades since the paper—are other motivators behind our project.

Language is, at its core, a system that is both digital and infinite. To my knowledge, there is no other biological system with these properties...—(Chomsky 1991).

Understanding a domain is defined as the ability to rapidly produce programs to deal with new problems as they arise in the domain.—(Baum 2009).

A computer ‘understands’ a subset of English if it accepts as input sentences from this subset and is capable of answering questions based on the information in the input.—(Bobrow 1964)

Self-consciousness, i.e. the ability to observe some of one’s own mental processes, is essential for full intelligence—(J. McCarthy 2008)

A cognitive system must think, improve by learning, adapt to the environment, and find structure—discover answers and insights to complex questions—in massive amounts of ambiguous, noisy real-world and domain knowledge. Such systems possess the ability to analyze a given problem from multiple perspectives and identify the viewpoint that synchronizes with the context; ascertain the problem objective, weigh multiple solution strategies and activate scheme(s) that can transport the system nearer to its goal; include commonsense reasoning (Lieberman et al. 2004; McCarthy 1959; Minsky 2000; Singh et al. 2004a, b) and improvise as well.

Thus, besides being a reason for contemplation on the fascinating abilities of our mental faculties, such a cognitive model of text comprehension could typically form the basis of ‘cognitive’ plagiarism detectors, library cataloguing systems and supports for children with reading disorders. The model also forms a platform for the merger of all the distributed research initiatives on the different aspects of language comprehension. Kowalski (2011) observes how human intelligence could benefit from computational thinking.

Our research, driven by curiosity, intuition and introspection, utilizes a polymath approach—drawing from psychology, philosophy, neuroscience and computer science—to work towards the solution. Psychology helps in understanding human nature, social and cultural influences on decisions, cognitive biases, etc.; Philosophy—to acquire knowledge on theories and questions on the mind, intelligence and thinking; Neuroscience—to appreciate the neural underpinnings of the human brain—a guide towards the abstraction of all that an artificial cognitive system needs to achieve; and Computer science leads to modeling the various elements of cognition, identified in the other sciences, and synthesis of requisite algorithms and architectures.

We do not claim to have excavated the answers to the questions posed at the beginning of the article, nor of having arrived at a complete model of text understanding that mimics the brain, but try to present a plausible scheme of the same. This article marks the first step of our attempts en route to understanding and emulating the processes leading to text comprehension in the human mind.

Our efforts meander through a top-down design process—a journey beginning at the macrocosm, driven towards the quark-view microcosm of ‘intelligent’ system design—and are roughly guided by the following steps:

- (a) Identification of the basic operations of the mind during text understanding.
- (b) Segregation of the operations into broad categories (or ‘agencies’).
- (c) Enumeration of the fine-grained ‘agents’ that underlie the agency-operations.
- (d) Construction of the elements of intra-agency and inter-agency communication and agent-activation.
- (e) Designing a model architecture that supports all of the above.

This paper focuses largely on steps (a) and (b) and provides a rough draft of elements that lead to (e), thus forming a blueprint in the nature of a requirements specification for our system design processes. We begin with an outline of the pre-requisites of a self-evolving cognitive system, followed by a list of the basic processes constituting text comprehension. This leads to discussions on the macro-components (mind-agencies and memory constructs) of the framework, the working principle and synthesis issues. The framework is analyzed through a dry-simulation and is corroborated by human subjects, a study of correspondences with the human brain and Minsky’s model of the mind, and conceptual comparisons with existing ‘cognitive’ ‘conscious’ architectures.

The novelty in our work lies in using Minsky’s model of the human mind to design a framework for cognitive language understanding. The system aims to formulate a bespoke procedure of comprehension that best fits a problem, learn from mistakes and improvise as well. While existing language understanders either do not ‘reflect’ or are not ‘self-reflective or ‘self-conscious’ or do not indicate the possession of intuition and commonsense, our framework includes each of these elements. The design is currently in its very early stages and is prone to evolution with our recurrent knowledge gain and clarification of concepts on the brain-processes.

The article begins with a brief introduction to the key inspirations underlying our concepts (Sect. 2), followed by the basics of the foundation theories (Sect. 3), a description of the proposed concepts (Sect. 4), and an analysis of the strengths of the framework (Sect. 5). It ends with a summary of the key ideas introduced herein and our future work directions (Sect. 6).

## 2 Related work

This section begins with a tribute to (Turing 1950), wherein the question ‘Can machines think?’ laid the foundations for artificial intelligence and its derivatives. Our investigations, motivated by Turing’s phenomenal article, aims to contribute to ‘thinking-machine’ research endeavors; perhaps lead to a methodology for the measurement of MIQ (Zadeh 1994) in terms of language comprehension.

Primarily based on Minsky’s theories on the ‘Society of Mind’ and the ‘Emotion Machine’, our work is influenced by and draws from pioneer research efforts on machine-

text understanding over the last four decades. The rest of this section chronologically introduces these projects.

Turing in (1949) describes the design of random unorganized self-organized structures for the construction of intelligent machines—built on the human model that begins as mechanisms with no capacity to handle elaborate operations, but through a gradual processing of interferences, develops mature handling capabilities. The pleasure-pain system outlined here is perhaps the earliest work on ‘understanders’ built using CWW (Zadeh 1996) to quantify and process degrees of ‘certainty’ (‘tentative’, ‘uncertain’ and ‘definite’) of pleasure and pain ‘affects’.

Bobrow (1964) describes a pioneering attempt towards defining natural language structures to capacitate the computer into solving algebraic problems in the form of stories. Winston (1970) adds to the concepts in Bobrow (1964), by concentrating on the construction of programs that empower a computer into forming and manipulating abstractions of a given scenario via visual-concept extraction skills.

SHRDLU in Winograd (1971), is one of the first and finest efforts at formulating computing mechanisms that ‘understand’ and communicate in English. The system uses syntax, semantic and deduction principles [based on Hewitt (1970)], and context to disambiguate senses, and uses procedures to represent knowledge. The system is thus able to activate knowledge instances on need and emulate comprehension through procedural forms. Charniak (1972) is a treatise on the development of a model for story comprehension by children. Besides focusing on the syntactic and the semantic elements, the work stresses on the incorporation of real-world knowledge, context and relevance-extraction towards comprehension.

The concept of the ‘Answer-Library’—an ever-growing performance library of procedures learnt or endogenously constructed, and indexed by problems for which the procedure was appropriate, in Sussman (1973), is a major inspiration in our design. The described model, ‘Hacker’ focuses on intellectual skill acquisition, within a domain of discourse; where given a situation, the system either recalls appropriate procedures, or in the worst case, writes procedures of its own; system performance improves with experience.

Sloman (1978) on ‘philosophical thinking and its transformation in the light of computing’, illustrates essential concepts on multi-perspective visualization of a situation and the layers of reflection of the human mind which laid down the foundations of the ‘CogAff architecture (Sloman 2001)’ of the mind. These perceptions find fine-grained extension in Minsky’s phenomenal compilation on the ‘Society of Mind’ theory (Minsky 1986)—the basic components of which are described in Sect. 3.2.1.

Minsky (1986) is the ultimate culmination of a computational theory of the human mind; not only is it a

collection of theories, but also a consequential catalyst for ‘thinking’ on ‘thinking’. The notions introduced here are extended in the ‘Emotion Machine’ concept (Minsky 2006), where the author presents a six-layered structure of the human mind and a computational theory for ‘thinking’ and ‘emotions’. It is indeed surprising that over the last three decades or so, since its inception, there have hardly been any notable initiatives towards the realization of Minsky’s theories. Some attempts are:

- (a) ‘DUAL’ (Kokinov 1989, 1994) which describes the integration of symbolic and connectionist architectures to form a cohort of small-scale agents that respond to changes in context and the environment. In its present state, the architecture does not incorporate the different realms of ‘thinking’ represented by the ‘critic-selector’ architecture in Minsky (1986), (2006), Singh et al. (2004), Singh (2003), Singh and Minsky (2003, 2004).
- (b) ‘EM-ONE (Singh 2005)’, a contemporary venture on the development of an emotion machine. It realizes the lower three layers of Minsky’s model of the human mind.
- (c) ‘FUNK2 (Morgan 2010)’ a programming language focusing on the emulation of efficient ‘meta-reasoning and procedural reflection’. Morgan (2013) extends the same towards the emulation of the four lower layers of Minsky’s model.

Adding to Minsky theories is McCarthy’s work on the emulation of commonsense reasoning (McCarthy 1959), and machine consciousness (McCarthy 1995, 2008). These reinforce the notions of the possession of real-world knowledge and consciousness as pre-requisites of an ‘intelligent’ system.

CMATTIE (McCauley et al. 2000; Zhang 1998), IDA (Baars 1988; Franklin 2003) and LIDA (Franklin and Patterson 2006; Sneider 2011) are some very recent pursuits towards the design of ‘conscious’ software agents that emote, reflect and learn and serve as frameworks for Artificial General Intelligence. These are based on the theories of the ‘Society of Mind’ and ‘blackboard architecture’ inspired global workspace (Baars 1988, 1997, 2002) theory.

Principles of ‘blackboard architecture (Erman et al. 1980)’ have largely influenced our model. This architecture is guided by the rigors of opportunist scheduling across a number of software specialist agents that ‘brainstorm’ over solutions to a problem. A ‘blackboard’ serves as a shared repository of agent-contributions towards problem-solving. Hayes-Roth (1985) describes the use of the architecture for the emulation of cognitive reflection. The advantages and disadvantages of the architecture are succinctly described in Hunt (2002). Our framework uses blackboard-type structures not only to list agency-suggestions but also as a

mechanism for the system to ‘reflect’ upon errors and ‘learn’ from them.

In Baum (2009) we find a conceptualization of the synthesis of flexible self-assembling programs, an ‘Artificial Genie’, that understands. The phenomenon of ‘understanding’ is to be brought about through agents or modules called by context-dependent causal domain simulation positions—such that the computations are meaningful in the real-world. The system is to include processes that mimic adaptation and consequent survival in a competitive environment; concise modules and code-scaffolds are to enhance the speed of execution.

As a last note, a mention of two present-day successful natural language understanders would help clarify the purpose of our research. Both MontyLingua (Liu 2004) and the DeepQA architecture (Ferrucci et al. 2010) [underlying Watson—the winner of Jeopardy! 2011] have displayed unparalleled success towards capacitating a machine to comprehend language; the former is robust, does not require training and is enriched in commonsense (Havasi 2007; Lieberman et al. 2004; Singh et al. 2004), while the latter, though lacking in commonsense, works in real-time and can compete with human beings. However, neither endorse ‘thinking’ or ‘reflecting’ (Grosz 2012; Liu 2004), and are thus far from being truly intelligent as was envisioned in Turing (1950). ‘Thinking’ across all the mind-layers in Minsky (2006) for language understanding is precisely what we wish to address.

With this brief description of the influences, the article now moves on to a discussion of the theories underlying the proposed computational mind-agency architecture.

### 3 Theory

This section begins with an overview of the brain activities underpinning language processing. This serves as our design guide—indicating all the processes that an artificial system requires to accomplish, if not imitate. This is followed by a discussion of the highlights of Minsky’s theories.

#### 3.1 Brain functions in language processing

When reading, the brain executes a deft series of intricate eye movements that scan and fixate within words to extract a series of lines and edge combinations (letters) forming intricate spatiotemporal patterns. These patterns serve as keys to unlock a tome of linguistic knowledge, bathing the brain in the sights, sounds, smells, and physicality of the words’ meaning. It is astounding that this complex

functionality is mediated by a small network of tightly connected, but spatially distant, brain areas. This gives hope that distinct brain functions may be supported by signature subnetworks throughout the brain that facilitate information flow, integration, and cooperation across functionally differentiated, distributed centers.—(Modha et al. 2011).

Reading is a cerebral activity concerned with the abstract—thoughts, ideas, tones, themes and metaphors. The human brain does not possess neural circuits dedicated to reading (Wolf 2007), but forms these circuits by weaving together different regions of neural tissue devoted to other abilities, like object recognition, spoken language, motor coordination and vision.

Studies (Price 2000; Ramachandran and Blakeslee 1999) have shown that the cerebral cortex is the primary language processing center. The cerebral cortex, responsible for unsupervised learning, directs the brain’s higher cognitive and emotional functions. It is divided into two almost symmetrical halves—the cerebral hemispheres—each made up of four lobes—and connected by the corpus callosum. The parietal, temporal, and occipital lobes—all located in the posterior part of the cortex—organize sensory information into a coherent perceptual model of our environment centered on our body image; the frontal lobe is involved in planning actions and movement and abstract thought. The association areas within these lobes integrate multi-modal sensory information and relate it to past experiences, after which the brain makes a decision and sends nerve impulses to the motor areas to respond. These areas work in sync to produce all forms of conscious experience including perception, emotion, thought and planning, as well as unconscious cognitive and emotional processes. Table 1 summarizes the language processing functions of the lobes, and Table 2 highlights the memory categories of the brain.

Besides the cerebral cortex, the cerebellum (Eccles et al. 1967) plays a role in the formation of procedural memories brought on by supervised learning. Turing (1949), refers to the cortex as an unorganized machine and the human brain to be uncannily similar to a universal machine, but with far greater capacities.

We do not aim to design components that mimic the neural activities of the brain areas, but rather to emulate the functions of these areas to form granules of comprehension—networks of hypergraphs of coherent associations across interdisciplinary knowledge elements. The above tables serve as requirements specifications—akin an SRS document—for our system design processes.

Neuromorphic processors (Mead 1990) are being prominently investigated under the ‘Brains in Silicon’ and

**Table 1** The lobes in the cerebral cortex of the human brain and their language processing mechanisms

Lobe	Lobe functions	Functions typical to text/language comprehension
Occipital	Processes visual information and passes its conclusions to the parietal and temporal lobes	Integrates visual information, giving meaning to what is seen by relating the current stimulus to past experiences and knowledge
Frontal	Assists in motor control and complex cognitive processes like attention, reasoning, judgment, decision making, problem solving, learning, reasoning and strategic thinking, social behavior and relating the present to the future. Forms the working-memory and the prospective memory (Winograd 1988)	Broca's area—resolution of syntax and morphology Defines the 'self'
Parietal	Assists in processing multimodal sensory information, spatial interpretation, attention, and language comprehension	Angular gyrus—language and number processing, spatial cognition, memory retrieval, attention mediation and understanding metaphors (Ramachandran and Hubbard 2001, 2003)
Temporal	Assists in auditory perception, language comprehension and visual recognition, storing new memories—facts (semantic), events (episodic), autobiographical memory (Conway and Pleydell-Pearce 2000), and recognition (familiarity + recollection) memory (Rugg and Yonelinas 2003)	Wernicke's area—resolution of semantics and word meanings Amygdala— affective processing and memory consolidation (refer to McGaugh (2004) for affective influences on memory) Hippocampus—storage and consolidation of memories from the short-term to the long-term semantic (factual) and episodic (event) memory, and spatial navigation Basal ganglia—reinforcement learning, procedural memory, priming and automatic behaviors or habits, eye movements (Hikosaka et al. 2000) and cognition (Stocco et al. 2010)

**Table 2** Categories of human memories

Memory	Description
Working	Deals with temporary representations of information about the task that the organism is currently engaged in
Episodic	Remembers details of specific events; predominantly contextual; these memories can last a life time; underlies the emotions and personal associations with the event
Semantic	Learns facts and relationships between facts; predominantly non-contextual; the basis of abstractions of the real world through cross-factual associations
Declarative	Made up of memories consciously or explicitly stored and recalled; is constituted by episodic and semantic memories.
Procedural	Made up of memories pertaining to implicit learning leading to automatic behaviors; are unconsciously recalled
Long-term	Encodes information semantically (Baddeley 1966); comprises of declarative and procedural memory elements
Short-term	Encodes information acoustically (Baddeley 1966); memories recalled for duration of the order of seconds without repetition (rehearsal); does not encompass manipulation or organization of memories—as is for the working-memory
Sensory	Memories of sensory stimulus to the sensory perceptors, after the stimulus has ceased; is of the order of milliseconds
Visual	Explicit memories pertaining to visual experiences
Olfactory	Explicit memories pertaining to olfactory experiences
Haptic	Explicit memories pertaining to tactile or haptic experiences
Taste	Explicit memories pertaining to experiences of taste
Auditory	Explicit memories pertaining to auditory experiences
Autobiographic	A subset of the episodic memory; deals exclusively with personal experiences
Retrospective	The action of remembering content of the past
Prospective	The action of 'remembering to remember'; memories activated in the future based on time or event cues

'SyNapse (Modha et al. 2011)' projects. These initiatives focus on the emulation of the brain's neural activities, computing efficiency, size and power usage, whereas, we wish to simulate the mind—the control mechanisms that spike the neurons in the processors; linking cognition to cellular mechanisms.

### 3.2 Minsky's theories

#### 3.2.1 Fundamentals of the 'Society of Mind' theory

One could say but little about 'mental states', if one imagined the Mind to be a single, unitary thing. But,

if we envision a mind (or brain) as composed of many partial autonomous ‘agents’—a ‘Society’ of smaller mind—then we can interpret ‘mental state’ and ‘partial mental state’ in terms of subsets of the states of the parts of the mind. To develop this idea, we will imagine first that this Mental Society works much like any human administrative organization. On the largest scale are gross ‘Divisions’ that specialize in such areas as sensory processing, language, long range planning and so forth. Within each Division are multitudes of subspecialists—call them ‘agents’ that embody smaller elements of an individual’s knowledge, skills and methods. No single one of these agents knows very much by itself, but recognizes certain configurations of a few associates and responds by altering its states.—(Minsky 1986).

A modular, hierarchical theory; the principal constituents of the ‘Society of Mind’ that apply to the constructs discussed herein, are:

*Agents* An *agent* represents the building blocks of a computational mind; a component of a cognitive process that is simple enough to ‘understand’. An agent is a generalized complex granule (Jankowski 2013) with inbuilt control mechanisms.

*Agency* Societies of *agents* that in totality perform functions more complex than any single agent.

*K-lines* An agent with the purpose of turning on a particular set of agents. *Nemes* and *nomes* are two general classes of *k-lines*—analogous to the data and control lines in system architecture, respectively.

*Nemes* Agents responsible for the representation of an idea (context) or a state of the mind. Examples of *nemes* are—

*Polynemes* Stimulate partial states within multiple agencies—as a result of learning from experience—where each agency focuses on the representation of a particular aspect of a thing and thereby connecting the same thing to a number of ideas;

*Micronemes* Bestow ‘global’ contextual signals to agencies all across the brain and handle subtle elements—those which cannot be crisply defined or lack specific terminology—of situations.

*Nomes* Agents that control the manipulation of representations and effect agencies in a predetermined manner. Examples of *nomes* are—

*Isonomes* Trigger the same uniform cognitive operation across a multitude of agencies, implying the application of the same idea across a number of many things at once;

*Pronomes* Control the attachment of terminals to frames and are typically associated with the short-term memory representation of a particular role (e.g. actor, cause, trajectory) of an element;

*Paranomes* Operate on agencies across multiple mental realms simultaneously with identical effects across all of them.

*Frames* Form of knowledge representation associated with representation of an event and all its associated properties and components through *frame-slots*.

*Difference-engines* Problem solvers based on the identification of the dissimilarities between the current state of the mind and some goal state.

*Censors* Restrain mental activity that precedes unproductive or dangerous actions.

*Suppressors* Suppress unproductive or dangerous actions.

*Protospecialists* Highly evolved *agencies* that yield initial behavioral solutions to basic problems like locomotion, defense mechanisms etc. These develop with time. This concept acknowledges Noam Chomsky’s views on language skills being ‘hardwired’ in children (Chomsky 1959).

#### *Types of Learning*

*Accumulating* Remember every experience as a separate case.

*Unframing* Find a general description for multiple examples.

*Transframing* Form an analogy or mapping between two representations.

*Reformulation* Find new schemes of representing existing knowledge.

*Predestined learning* Learning that develops under sufficient internal and external constraints such that the goal is assured, like learning a language or learning to walk.

*Learning from attachment figures* Learning how and when to adopt a particular goal and prioritize it, based on reinforcement of knowledge by ‘attachment figures’—people who have an impact on our minds. E.g., ‘praise’ and ‘censure’ from parents and teachers contribute significantly to goal learning.

### 3.2.2 ‘Frames’ to represent knowledge

A ‘frame (Minsky 1975)’ is a data structure for representing typecast situations or events. It depicts a unit of information selected from memory, when one needs to store facts about a new encounter or if an existing perspective undergoes a major upheaval. It thus, reflects the

subjective time-sensitive view of a situation. Frames contain various types of information—specific data cues on a situation, information about how to use a frame, what (might) happens next, actions that may be taken if the expectations are not confirmed, etc.

These constructs form hierarchical connected graphs of nodes and relations, where ‘top-level’ frames carry a fixed abstraction of the situation, while the ‘lower-level’ frames have terminal slots (which again are smaller frames or ‘sub-frames’) to carry specific data instances. The data entry into the terminals is guided by assignment conditions like ‘name of a person’, ‘pointer to another sub-frame’, ‘relation to another sub-frame’, etc. Collections of related frames form frame-systems, where effects of important actions are mirrored by transformations across frames in a system and each frame might represent a different perspective of the current situation.

A frame-system is activated by an information retrieval network that detects frames as situation-representatives and correspondingly initiates matching algorithms to assign values to the frame’s terminals, consistent with the context-sensitive assignment-conditions, system expectations or surprises and the envisioned system goal.

In language, syntactic structural rules and semantics direct the selection and assembly of transient sentence frames. These frames are predictably complex structures—requiring the appropriate encoding of textual temporal and spatial elements to allow causal frame transformations. The basic frame-types for representation of linguistic entities are as follows, and understandably, these denote different levels of comprehension-granularity:

*Surface syntactic frames* For verb and noun structures, prepositional and word-order indicator conventions.

*Surface semantic frames* For action-centered meanings of words, qualifiers and relations involving participants, instruments, trajectories and strategies, goals, consequences and side-effects.

*Thematic frames* For scenarios concerned with topics, activities, portraits, setting, outstanding problems and strategies commonly connected with a topic.

*Narrative frames* For skeleton forms for typical stories, explanations, and arguments, conventions about foci, protagonists, plot forms, development, etc.; designed to help a reader or a listener construct a new, instantiated *thematic frame* in the mind.

Intuitively, every word ( $x$ ) in the human lexicon exists in the memory in all the three forms of the frame-topology, i.e., frames, terminals and slots. The nature of activation of facts associated with  $x$  in the memory depends on  $x$ ’s role in the context being processed. For example, in the sentence, ‘Jane loves spring’—the word ‘spring’ leads to the activation of a frame of the same name; while in the

sentence, ‘Jane is an all-season person’—the word ‘spring’ crops up in the memory in its capacity as a terminal or a slot.

### 3.2.3 Thinking

A model of language-understanding cannot be ‘cognitive’ if it does not ‘think’. ‘Thinking’ stands for a complex phenomenon entailing the analysis of a given situation across a number of causal perspectives, consideration of valid propositions and solution prescriptions, and to apply or improvise upon them towards the appropriate solution. This involves processes of recall, manipulation and organization of a vast repertoire of real-world and domain knowledge, and far richer automated reasoning processes than those known in AI, i.e., a meta-theory of reasoning. Human ‘thinking’ operates across a diverse array of mental realms (Singh and Minsky 2003, 2004), some of which are—

*Physical* Where object behavior is predicted;  
*Social* Dealing with inter-personal relationships; and,  
*Mental* Reflections upon mistakes, failures and successes.

In Minsky (2006), ‘thinking’ is envisioned in terms of ‘critic-selector’ (Singh et al. 2004; Singh 2003; Singh and Minsky 2003, 2004) model of the human mind—a representation of ‘reflective thinking’. The keynote of this model is, given a problem, instead of applying a particular general-purpose method for inference or action, the system analyzes (‘criticizes’) its knowledge of AI techniques to choose (‘select’) the one that is best suited to the problem (analogous to the causal-diversity matrix in Minsky (1992)). In other words, the system ‘thinks’ briefly on how it should ‘think’ about the given problem, and then ‘thinks’ about it as per the chosen method.

The six-layered architecture in Fig. 1 depicts his model of the human mind. Each of the layers incorporates ‘critics’ that assess the situation in the external world as well as the internal system states and activate ‘selectors’ that accordingly initiate ‘thinking’ on the interpretation strategies. The lower levels of the model handle and represent ‘instinctive reactions’ to the external world, while the higher levels control the reactions of the lower levels in accordance with the system’s model of itself. These layers symbolize multi-realm ‘thinking’. Figure 2, is a pictorial representation of the functioning of the critics and the selectors across the lower three layers of the mind-model.

The basic functions of the layers in the model have been defined as follows (Minsky 2006; Singh and Minsky 2004):

An average human being is born with instincts that aid survival—an implicit database of ‘if situation and goal, then do action’ reaction-rules like: ‘if there is a seat and

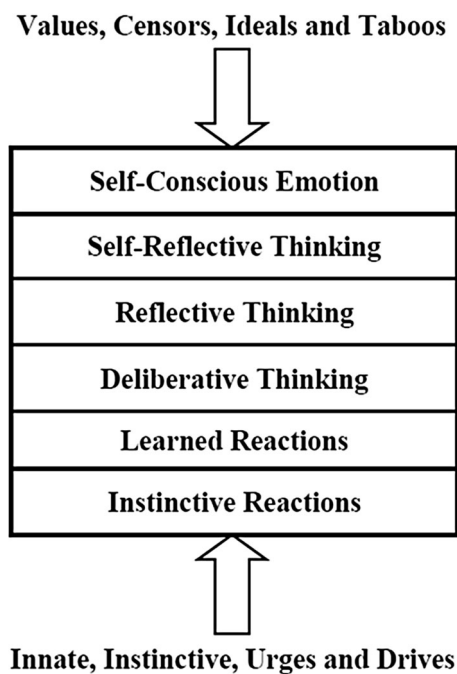


Fig. 1 The six-layered model of the mind (Minsky 2006)

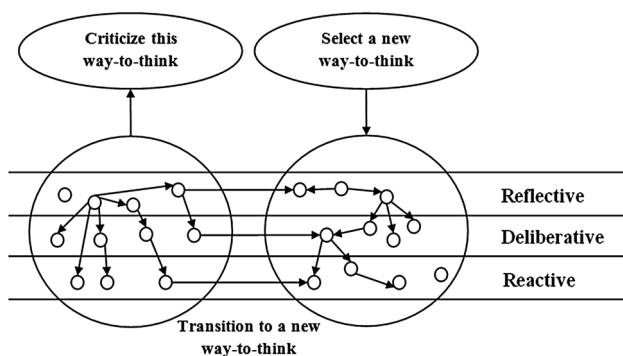


Fig. 2 A ‘critic-selector’ model of thinking. The small circles represent agents and other resources specific to that way-to-think, spanning the many levels of the architecture (Singh and Minsky 2003)

you are tired sit’. Such rules are often instrumental in predicting outcomes to situations. E.g. I am far from something I want → Move towards it; I feel scared → Run away.

**Learned reactions** Life teaches one that certain conditions need specific ways of being handled, thereby creating a ‘learned reactions’ database of <problem\_descriptors, action, result, reason> tuples ranked in the decreasing order of reinforcement; greater the reinforcement, higher is the probability of the action being recalled. E.g I am far from something I want immediately → Run towards it; I feel scared → Run quickly to a safe place.

**Deliberative thinking** Consideration of several alternative solution approaches, and choosing the best; using

logic and commonsense reasoning to select solution paths. E.g Action A did not quite achieve my goal → Try harder, or try to find out why; Action A worked but had bad side effects → Try some variant of that action; Achieving goal X made goal Y harder → Try them in the opposite order.

**Reflective thinking** Introspection over the mental activities that went into arriving at the decision, rank inference methods, representation selection, etc. E.g.: The search has become too extensive → Find methods that yield fewer alternatives; Overlooked some critical feature → Revise the problem description; Cannot decide which strategy to use → Formulate this as a new problem.

**Self-reflective thinking** Reflection on oneself as a ‘thinker’. While the reflective layer considers only recent thoughts that went into some decision-making, the self-reflective layer focuses on the entity that ‘thought’. E.g I missed an opportunity by not acting quickly enough → Set up a mental alarm that warns me whenever I am about to do that; I can never get this exactly right → Spend more time practicing that skill.

**Self-conscious emotion** Verification of accordance of decisions with ideals, include self-appraisal by comparing one’s abilities with others. E.g I think I am good at this task → Can I do it as well as the best people I know?; My mentor would not have made this mistake → What would he have done in this situation?; How is it that other people can solve this problem? → Find someone good at this problem and spend time with them.

Following the definitions of the different levels of ‘thinking’ undertaken by the layers of the mind, the ‘critics’ and the ‘selectors’ in these layers require to lead to the following operations—with respect to text comprehension:

**Instinctive or inborn reactions** ‘Looking at text’—accept text inputs.

**Learned reactions** Assign meaning to the elements seen—alphabets, digits, special symbols, white-spaces, punctuation; agglomeration of symbols into words, numbers, codes, phrases, clauses, sentences; syntax and semantic analysis of the text extracted; literature categorization into prose, poem, etc.; genre resolution.

**Deliberative thinking** Disambiguation of word-meanings, sentence-meanings, genres; rhetoric and prosodic analysis; analyze relevance and coherence of flow of concepts across text; consolidate individual text-elements into concepts; visualize scenes.

**Reflective thinking** Reason and optimize deliberative thinking processes; generate curiosity (questions in the computational mind) and activate schemes to gratify the same; build cross-text and cross-contextual associations.

*Self-reflective thinking* Evaluate interest and comprehension progression through text; overcome cognitive biases and reform concepts; text section identification—introduction, rising action, climax, denouement and conclusion; regulate eye-tracking (re-read sections, monitor reading speed).

*Self-conscious emotion* Attachment of emotions or levels of interest and perceptions to the entire text; to what extent does the text come-up to the reader's expectations and ideals—is it taboo, inspirational, fun, tragic, unput-downable, etc.; will the reader recommend it to anyone; will the reader read it again; how does the current reading affect the reader—did the reader gain new knowledge, which concepts were clarified.

Clearly, the functions of the layers overlap (e.g. most functions under *learned reactions*, like assignment of symbol meaning arising out of commonsense or instinctively post learning-reinforcement over a sufficiently long time-frame; *deliberative, reflective, self-reflective* and *self-conscious* thinking are concurrent co-operative processes) and information percolates in the bottom-up as well as the top-down directions. The information that is transferred to the higher layers relies on the extracted text-sample while that from the higher layers is conceptual and relates to the reader's sensibilities—acquired through learning, experience and commonsense reasoning.

The layers involved in the generation and manipulations of the frames are in the following order:

*Surface syntactic frames* Instinctive, learned and deliberative thinking.

*Surface semantic frames* Instinctive, learned, deliberative and reflective thinking.

*Thematic frames* Deliberative, reflective and self-reflective thinking, and self-conscious emotion.

*Narrative frames* Deliberative, reflective and self-reflective thinking, and self-conscious emotion.

Hereon, the article proceeds towards an elucidation of the proposed concept—an illustration of the design requirements of an intelligent system, followed by an outline of the basic processes of text comprehension which leads to the enumeration of the components of the framework, its working principle and issues particular to its realization.

#### 4 The proposed framework—design and synthesis of a computational mind

This section is dedicated to a description of the intended agency-architecture for machine understanding, focusing particularly on the phenomenon of text comprehension. The description begins with a brief study on the essentials

of a self-evolving computational system, and an abstraction of the tasks that the mind performs during language comprehension, thus laying the foundations of our design initiatives.

The study leads to the explication of the conceptual framework of the computational mind-agency architecture—an elucidation on the mind-agencies (functions and interactivity) and memory constructs, and related synthesis issues.

##### 4.1 Designing a self-evolving computational system

You end up with a tremendous respect for a human being if you're a roboticist—Joseph Engelberger, 1985

The human mind is a continuously evolving computational system that acquires, builds, stores and manipulates symbols; an infinite (countably infinite?) state machine to be precise. Thus, the emulation of the mind towards the construction of a 'thinking-machine' calls for the reduction of an infinite-state machine to a finite-state one. A 'very hard' problem undoubtedly, but nonetheless an opportunity for scientific analysis of the questions asked by mind-philosophers, introspection and observation on the mind-processes, and defining heuristics towards its emulation.

Drawing from the concepts in Backus (1978), Erman et al. (1980), Harrison and Minsky (1992), Sloman (1984) the design prerogatives of a naturally evolving computing system, akin to the human brain, can be summarized into the following points:

- (a) Possess a finite alphabet set—primitive language elements which can be modeled into complex components like words etc.
- (b) Have a substantial, yet finite, memory unit that can store a large number of independently variable symbols. The symbols assume values from elements in the alphabet set, and the cardinality of valid symbols and that of the alphabet set dictate the number of states that the system can be in.
  - (a) Values of the symbols may represent data or instructions.
  - (b) These values can be generated, stored, searched for, manipulated upon and deleted, implying that the system includes a large and adaptive repository of information or knowledge.
  - (c) Knowledge includes intuition and commonsense as well as run-time concepts (partial, complete, correct and incorrect) generated in the process 'understanding' the real-world.
  - (d) Mechanisms to handle knowledge include strategies to associate between cross-domain knowledge (Bush 1945), divide knowledge

- into context-sensitive units and to use them selectively and efficiently. Choices for these design issues need to exploit sources of structure and constraints intrinsic to the problem domain.
- (e) Symbols interpreted as instructions should control the internal and the external behavior of the system, generate behavior, exhibit self-control as well as be self-modifying. Some of these instructions are to be conditional—typically underpinning adaptable and intelligent behavior, and learning based on environmental influences and feedback (positive and negative).
  - (f) Some of the symbols may represent the information flowing into the system through sensors and other input devices, and can be used by conditional instructions. The system can thus treat its symbols as representatives of beliefs about the world.
  - (g) Besides primitive symbolic instructions which directly cause processes to occur, the use of symbols with meaning allows instructions, like assertions, to refer to an external world and be goal-directed.
- (c) An adaptive system requires being reflective (Maes 1987) or history-sensitive (Backus 1978) and self-conscious (McCarthy 2008), i.e., incorporate structures representing (aspects of) itself, allowing the system to question its own actions, answer and improve towards robustness and fault-tolerance. These include maintaining performance statistics for debugging (Ashby 1952), stepping and tracing facilities, interfacing with the external world, computation about future computations (or reasoning about control), self-optimisation, self-modification and self-activation.
  - (d) Require structures that represent the properties of the environment—complexity, variety, unpredictability and degrees of familiarity. This further imposes constraints on the types of perceptual systems required, kinds of belief representations, planning and executing mechanisms, learning mechanisms, etc.
  - (e) Emulate neurogenesis (Chugani et al. 2001) by being part of a social system—be able to acquire new forms of knowledge (e.g. new concepts, new languages and language skills) and be capable of adapting to various kinds of changes, modify some of their rules of behavior
- to cope with changing social needs, draw lessons from situations, differentiate between right and wrong (following established social norms), act unselfishly, recognize emotions and mood variances and react accordingly, identify levels of social hierarchy, etc.
- (f) The need to cope with a relatively large number of changing goals, principles, ideals, preferences, likes, dislikes—not all mutually commensurable or simultaneously compliable. This implies a need for motive-comparators ['critics and selectors' (Minsky 2006; Singh 2003; Singh and Minsky 2003)] and strategies for deciding between incommensurable alternatives, decisions based on long-term or short term objectives, and the ability to ignore or suppress some motives or needs in the light of others and form new goals.
  - (g) The system must be comparable or even faster than average human processing (Baars 1988)—conscious processing (of the order of 100 ms) and unconscious processing (at the speed of neural firing which is 40–1,000 times per second).

#### 4.2 Basic text comprehension operations and the layers of thinking

Assuming the different units of language like words, phrases and sentences are extracted, and that the text being processed is devoid of non-alphabetic elements (pictures and diagrams); comprehension involves a complex plethora of conscious and omniscient unconscious cognitive processes that ideally lead to the following mind-activities:

*Prediction* Envisage a future action—involving causally relating the present to past experiences and judging expectations on the basis of intuition, commonsense, reinforced learning and reflection.

*Visualization* Conjure mind-images [real or intentional (Husserl 1970)] of text components (people, places, events).

*Connection* Build factual or conceptual associations between: (a) frames recalled and those created for the current text processing event, and (b) existing real-world or domain knowledge and new information.

*Question and clarification* Reason (reflect upon), test the strength, completeness, correctness and relevance of constructed knowledge associations, leading to re-organization or rectification of the associations.

*Evaluation* Test the coherence between the perception granules, measure relevance of each and prune the

insignificant; attach notions of subjectivity or 'self-consciousness' (emotions, degrees of interest, summarize, biases, etc.) to the text as a whole as well as the constituent components.

Intuitively,

- (a) *Prediction* and *visualization* involves all but the topmost two layers of thinking; *connection*—the four lower layers; *question* and *clarification*—the learned, deliberative and reflective thinking layers; and *evaluation* involves the topmost three layers.
- (b) Reading and subsequent compression iterates (Ariely 2008) through the above stages—working incrementally on micro-granules of information to form coherent networks of information and a macro-granule summary of the text being 'read'.

The processes that underlie the above complex functions can be roughly outlined, in no specific order, as:

*Symbol-extraction and symbol-interpretation* Differentiation between foreground and background elements of the text-sample page, adjudge symbol boundaries, resolve ambiguities and stray markings; identification of the symbols as digits, alphabets, special characters, etc.

*Symbol-granulation* Group symbols into language granules—words, numbers, phrases, clauses, sentences, etc.

*Syntax-resolution* Identification of the syntactic nature (part of speech) of the symbol-granules.

*Semantic-resolution* Context-sensitive interpretation of the syntactic elements (words in general); involves intuitive and commonsense reasoning, deliberation and reflection over interpretations; support 'on the fly' interpretations of unfamiliar words and phrases from surrounding text and the genre. These further involve—

*Anaphora/cataphora-resolution* Resolution of the dependencies between explicitly and implicitly stated object-pronoun and person-pronoun elements.

*Spatio-temporal\_sense-resolution* Resolution of the temporal and spatial meanings of prepositional words or phrases.

*Context-resolution* Identification of the discourse-context and the text-genre.

*Sense-resolution* Identification of the correct context-sensitive meaning of homonymous words or phrases; resolution of the figure of speech of text elements.

*Relevance-evaluation* Identification of the importance of the words/phrases extracted and 'understood'; pruning away insignificant or un-required frame-elements; leads to summarization.

*Affect-evaluation* Monitor the progression of interest and affects across the text; identification of text sections—

introduction, rising action, climax, denouement and conclusion; assign affects to characters and sections.

*Comprehension-evaluation* Evaluation of the correctness, completeness and strength of comprehension; initiation of 're-reading sections' or modulation of reading speed according to the degree of comprehension and interest.

*Frame-generation/retrieval/manipulation* Creation, recall and operate upon frames and frame-systems to form concept granules (Jankowski 2013) across different level of granularity (syntax, semantic, narrative, thematic).

*Encoding/decoding* Translation of frames and frame-systems into suitably compressed, indexed and customized (flavored by parameters of 'self-consciousness') knowledge components, and vice versa; seamless integration of data-types (visual, audio, auditory, etc.) representing the same memory.

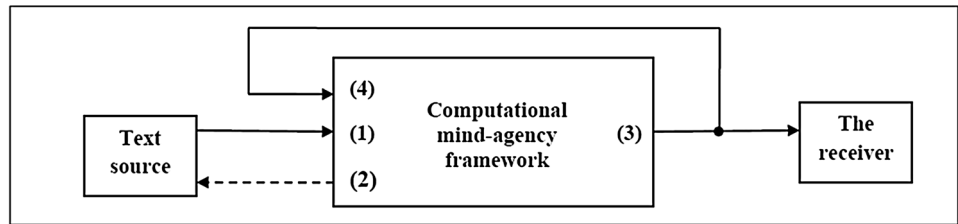
*Memory-handling* Short-term sensory information handling for symbol extraction/interpretation/granulation; declarative or procedural experience retrieval; activation of sensory experiences to effect affectual responses; short-term to long-term information consolidation; working-memory handling—monitor working sets of frames.

*Error-handling* Disambiguation of incorrect, unexpected or incomplete symbols or syntactic elements; suppress incorrectly activated word senses and contexts, consequently activate the correct senses, and propagate rectifications across currently active frames to update comprehension; update incorrect instances of existing knowledge and associated affects; overcome errors due to cognitive biases (Ariely 2008; Banaji and Greenwald 2013).

Instinctively:

- (a) These processes are complex, mostly concurrent, and co-operative, as has been hypothetically envisioned in Sect. 4.3.3 and depicted in Fig. 8.
- (b) Text comprehension ideally follows an 'iterative-incremental development (Ariely 2008)' execution scheme through the above processes (Fig. 3 is an abstraction of the scheme—the components of the computations mind-agency framework are elucidated in Sect. 4.3).
- (c) The 'meaning' of a word or a phrase implies the manner in which the sense of the language unit is encoded in the mind. These encodings could be in the form of precise codes in the native language of the system or as metaphors, synonyms or associations with other words. A single word or phrase may have multiple sensory (visual, auditory, etc.) implications (as shown in Table 2) as well.

**Fig. 3** An abstraction of the iterative-incremental-developmental strategy of comprehension. (**Vision** and **Deducer** are components of the mind-agency framework and imply the ‘eyes’ and the ‘brain’ of the system, respectively. Section 4.3 describes these components in detail)



**Index –**

- (1) – Input to the system = the text-elements.
- (2) – **Vision** might be instructed, by the **Deducer**, to re-read sections of the current text-source.
- (3) – Output of the system = a granule of comprehension = deep-semantics (cross-contextual information associations) + reflections + emotions + new-knowledge + updated old-knowledge.
- (4) – Granules of comprehension, produced at time ( $t_i$ ), are re-introduced into the system along with the current text-inputs at time ( $t_{i+1}$ ). The elements along the feed-back path influence the current comprehension – set the context of interpretation, facilitate anaphora/cataphora resolution, etc.

At time ( $t_1$ ), when the input to the system is the first text-element of the current session, values in the feedback path symbolize the state of the computational mind – cognitive biases arising out of comprehension processes handled in the recent past. Ideally, the system is robust and is able to override these exogenous influences during the course of current processing activities.

- (d) *Symbol-extraction/interpretation* involves the two bottom layers of thinking; *symbol-granulation*—the learned reactions layer; the remaining processes engage all the layers of thinking.
- (e) *Frame-generation/retrieval/manipulation, encoding/decoding, memory-handling, error-handling* are processes that support each of those preceding them in the above list.
- (f) The functions straddle multiple layers of thinking and require bi-directional information percolation. The information that is transferred to the higher layers relies on the extracted text-sample while that from the higher layers is conceptual and relates to the reader’s sensibilities acquired through learning, experience and commonsense reasoning.
- (g) These processes not only apply to text comprehension, but also to understanding in general—where instead of text, the computational mind processes simultaneous multi-modal sensory inputs from the environment it is in.
- (h) This list cannot be an exhaustive enumeration of the broad mechanisms leading to comprehension, and we strive to add to it as we recurrently enrich ourselves with the knowledge of the way the brain ‘understands’ the real world.

4.3 The computational mind-agency framework for text comprehension

Our brain is not a hierarchical control system. It’s more like anarchy with some elements of democracy.—(Dennett 2013)

This section focuses on the description of the macrocosmic elements of the computational mind-architecture—the mind-agencies, long-term memory databases of knowledge and the working-memory constructs; followed by an elucidation of the working principle of the framework and implementation issues. Elaborations on agent-structures and algorithms, and detailed memory data structure formats, though are out of the scope of this article, are our future research pursuits.

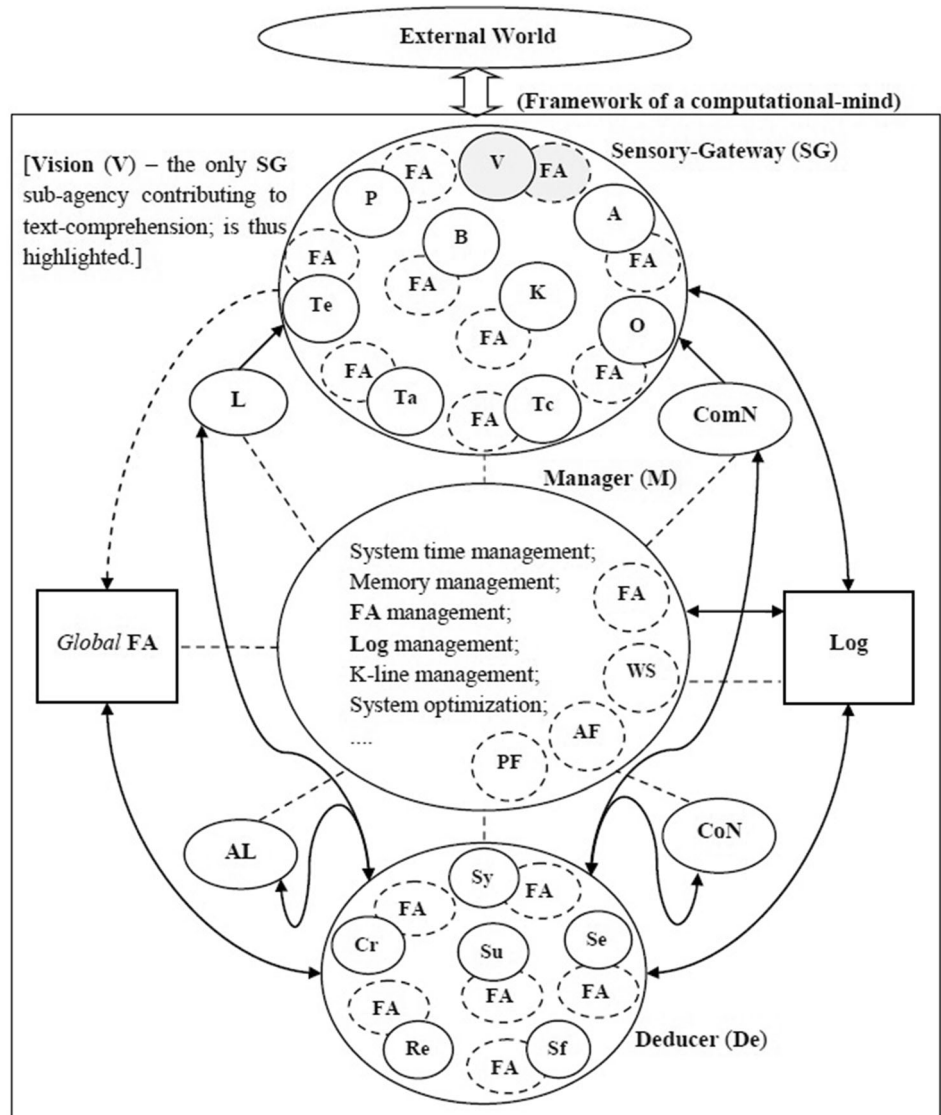
A computational mind is typically able to co-operatively process concurrent multi-modal sensory inputs harmoniously with existing knowledge about the real-world and the problem domain. Accordingly, each of the agencies enumerated here have multiple functions towards the realization of mind-processes. Our focus, however, being entirely on the text understanding processes of the mind, the framework components have only their roles towards text comprehension elucidated here.

4.3.1 Components of the framework

We have categorized the mind-agencies into **super-agencies**, each of which denote a complex cognitive functionality like ‘reasoning’ or ‘processing’, and **sub-agencies**. A super-agency comprises of a cluster of sub-agencies, each of which realize an operation that lead to the super-agency functionality. The sub-agencies are again built of agents, where an agent represents an atomic process underlying a sub-agency operation. Figure 4 is a pictorial representation of the mind-agency framework.

The super-agencies and constituent sub-agencies of text-comprehension in a computational mind are:

**Fig. 4** The computational mind-agency framework for text comprehension, and understanding in general



**Index –**

- ↔ Read and Write operations
- Read-only operation
- - - -> Write-only operation
- · · · · Control operations

(Arrow-direction indicates operation destination)

V	Vision	P	Pain	WS	Working-Set
A	Audition	Sy	Syntax	AF	Active-Frames
O	Olfaction	Se	Semantic	PF	Passive-Frames
Tc	Tactile	Sf	Self	L	Lexicon
Ta	Taste	Re	Recall	ComN	Commonsense-Network
Te	Temperature	Cr	Creative	CoN	Concept-Network
K	Kinesthetic	Su	Summary	AL	Answer-Library
B	Balance	FA	Frame-Associations	FA	Local FA

(a) *Sensory\_gateway (SG)* At any instant, SG serves as the receiver of sensory information, whereupon depending on the nature of the sensory-input, dedicated ‘sensory’ sub-agencies [**Vision (V)**, **Audition (A)**, **Olfaction (O)**, **Tactile (Tc)**, **Taste (Ta)**, **Balance (B)** (Robinson and Aronica 2013),

**Temperature (Te)** (Robinson and Aronica 2013), **Pain (P)** (Robinson and Aronica 2013) and **Kinesthetic (K)** (Robinson and Aronica 2013)] activate other framework components for further processing. SG transports system results to the external world as well.

Sub-agencies like **A**, **O** and **Te** continually receive stimuli from the environment and process these unconsciously; **Tc** and **K** are activated in the 'turning pages', 'scrolling over text' activities. However, none of these contribute significantly to the 'text comprehension' phenomenon and have thus not been elaborated upon. Our concern, here, being the synthesis of a computational mind towards text comprehension, the functions of the **Vision** sub-agency is where our interest lies.

- (1) *Vision (V)* The 'eyes' of the system—leads to textual symbol-extraction, symbol-interpretation and symbol-granulation.
- (b) *Deducer (De)* The 'brain' of the system; is responsible for all the text processing and comprehension activities. It receives outputs (data) of **SG** to formulate units (frames) of comprehension—utilizing syntax and semantic analysis mechanisms, relevance-evaluation, affect-evaluation, comprehension-evaluation and error-handling processes; sends out instructions (activation, re-evaluation, error signals, inhibition) to the other super-agencies as well. The sub-agencies of interest are:
  - (a) *Syntax (Sy)* Is responsible for syntax-resolution of the text-unit being processed and consequent generation and manipulation of surface syntactic frames.
  - (b) *Semantic (Se)* Is responsible for the identification of the literature category and text-genre, semantic-resolution of the text unit being processed in the light of the genre-context, and generation and updating of surface semantic, narrative and thematic frames.
  - (c) *Self (Sf)* Is responsible for seasoning all comprehension granules with values that define the system personality, i.e., introducing subjectivity (immune from cognitive biases) into text processing; multiple mental-realm activations.
  - (d) *Recall (Re)* Is responsible for thin-slicing a problem into sub-problems, mapping problems to memories and retrieving the same from long-term memory for processing in the current context.
  - (e) *Creative (Cr)* Is responsible for projecting and suggesting solutions for problems with no prior experience; the hub of reflection, imagination, creativity and system IQ.
  - (f) *Summary (Su)* Is responsible for analyzing the distance between the current state of the system and the projected goal through relevance, affect and comprehension progression

evaluation; can activate or inhibit agencies (under **De** and **SG**) based on summary results; consolidation of memories, both current and past.

- (c) *Manager (M)* The global administrator or 'heart' of the system; it runs in the background and is responsible for the activation and execution of 'involuntary' functions (system-time management, memory handling, process synchronization, K-line management, frame encoding/decoding, job scheduling, etc.) that support the functioning of all the other agencies; continual self-evaluation of system processes and subsequent updating towards improved (cost effective and robust) system performance.

The sub-agencies under **M** have not been elucidated as the functions thereof are typical system operations unparticular to text comprehension.

The databases—long-term memory stores of knowledge, that support the functioning of the agencies, can be enumerated as follows:

- (a) *Lexicon (L)* The vocabulary of the system; a resource of language units—words, phrases, idioms—and their meanings encoded in machine 'understandable' form; includes meanings of words 'learnt on the fly' and jargon; the meanings may be encoded into precise statements as well as exist in a number of data types (sounds, images, metaphors)—indicating the different ways the machine 'understands' or 'remembers' an element.
- (b) *Answer-library (AL)* A resource of  $\langle \text{solution\_strategy, result, reasons} \rangle$  for a given  $\langle \text{context\_parameters, problem} \rangle$  query.
- (c) *Concept-network (CoN)* Network of networks of inter-contextual concept granules, a hypergraph of associations across frame-systems; elements are retrieved 'consciously'.
- (d) *Commonsense-network (ComN)* Network of networks of commonsense and intuitive (automatic) behaviors; is the root of all information retrieval, i.e., the elements are retrieved 'unconsciously'; elements of **L**, **CoN** and **AL** are incorporated into the **ComN** after prolonged periods of reinforcement.

The basic global working-memory data-structures are as follows; these are referenced by all the agencies and form the basis of deliberative and reflective actions of the system:

- (a) *Log* A blackboard or scratch pad where time-stamped entries of agency-activities are made; indicates the state of the system at any given instant, analyzing which—a number of agencies may be self-

activated, and the **De** might activate or inhibit agency functions, initiate mechanisms like intelligent backtracking (Stallman and Sussman 1977), generate error signals, etc.; serves as an indicator of solution strategy results and reasons thereof for the system to ‘reflect’ upon.

- (b) *Frame-associations (FA)* A blackboard or scratchpad for frame-system manipulations during the process of text ‘understanding’; comes in *global* and *local* (per sub-agency) categories; all frame recollections are placed in the *global FA* space, while sections of the *global FA* are copied into *local FA* for deliberations by sub-agencies; the *local FA* of the sub-agencies under **SG** is analogous to the sensory memory concept in the human brain; the sub-agencies under **De** use their *local FA* workspace to reason through the applicability of multiple solution-perspectives before globally ‘advocating (a <problem, solution, reason> tuple)’ frame manipulation processes through **Log**; the sub-agencies under **M**, use their *local FA* to reason through system optimization mechanisms that would best support some globally approved frame manipulation exercise; each sub-agency can share sections or all of its *local FA* with the other agencies; globally approved suggestions (by **Su**) are implemented in the *global FA* and all updates to existing networks of information, are reflected across the long-term memory networks; all local trials are annotated in *local FA* but the trial-results are annotated in **Log** and *global FA* for deliberation and reflection by the other agencies.

The system memory-management constructs, used by **M**, are:

- (a) *Working-set (WS)* Set of pointers to frame-networks in **FA** being referenced within a narrow time-window (intuitively, of the order of seconds).
- (b) *Active-frames (AF)* Set of pointers to frame-networks in **FA** being referenced within a broad time-window (intuitively of the order of minutes); **WS** is a subset of **AF**.
- (c) *Passive-frames (PF)* Set of pointers to frame-networks in **FA** that were members of **AF** but were pruned away due to insignificance or lack of use; instead of consolidating them back to the long-term memory, these frames remain available during the entire span of the processing of the current text for quick “on-demand” placement into **FA** for re-processing.

Observations:

- (a) Considering that the design of the framework is prone to evolution, as we gain knowledge about the processes that lead to the human brain behaving the

way it does, the primary advantage that agencies assigned with dedicated responsibilities provides is the ease with which an agency may be upgraded without affecting the design of the entire framework; introduction of new agencies or framework components would however require changes percolating across every level of the design. Figure 5 summarizes the nested-modular nature of the computational mind structure.

- (b) Distributed processing across the agencies is the key functional principle of the system. Each of these agencies implies a granule of control or operation stack.
- (c) The agencies are interconnected such that it forms a causal system. This is roughly demonstrated in the feed-back schematic of the system in Fig. 3.
- (d) **SG** depicts instinctive and learned behavior, while all the other agencies transverse all the layers of thinking.
- (e) **V** references **L** and **ComN**, and **De** references **CoN**, **ComN** and **AL**.
- (f) **CoN** and **ComN** are inspired by the basics of ‘ConceptNet (Havasi 2007)’, while **AL** is influenced by ‘Hacker (Sussman 1973)’.
- (g) Information storage and retrieval from each of these long-term knowledge databases involves encoding/decoding processes across frame-types and data-types.
- (h) **Log** being the basis of inter-agency communication, these costs are grossly reduced—any message on **Log** is equivalent to broadcasting it across all the agencies for reflection or deliberation.
- (i) **M** is responsible for arbitrating multiple log-access requests from a number of agencies; this calls for

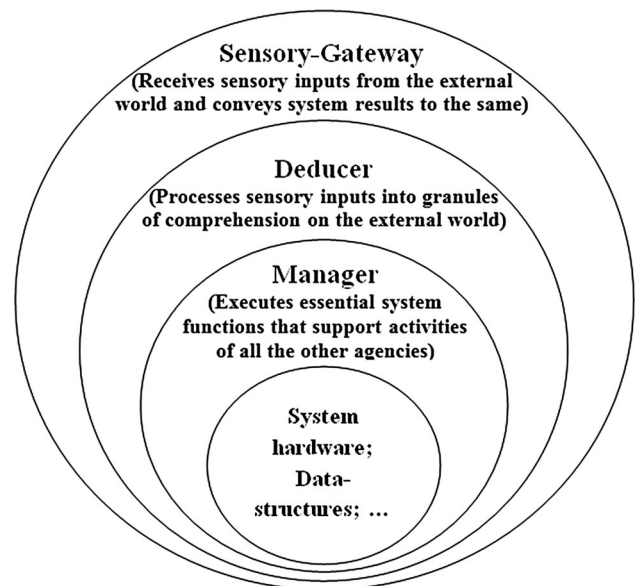


Fig. 5 A diagrammatic summary of the computational mind

- standard formats for **Log**-messages ('suggestions', 'applied methods', 'outputs', 'requests', etc.) for uniform comprehension across the system.
- (j) Following Minsky's terminology, **Re**, **Cr**, and **Su** form the *Difference-Engines* and **Su** the *Censor* and *Suppressor* of the framework.
- (k) **Su** is the control shell of the architecture—coordinating the inter-agency activity via heuristics and approximation schemes, to handle combinatorial explosions of thoughts and solution strategies, to ensure tractability of the text comprehension problem.
- (l) The sub-agencies under **De** can be categorized into the following, based on the levels of information-granules they deal with:
- Tier 1* Acknowledge system 'self'; subjective decisions—**Sf**.
  - Tier 2* Conjecture abstract or well-defined procedures for text interpretation—**Re**, **Cr**, **Su**.
  - Tier 3* Hypothesize steps of abstract procedures; procedure-step execution—**Se**, **Sy**.
- (m) *Global FA* and **Log** apparently resemble the global workspace (Baars 1988, 1997, 2002) construct of blackboard architectures (Erman et al. 1980). While the former is a platform for the formulation of frame-associations through agency operations; **Su** through standard **Log** message formats broadcasts the current status of the interpretation, through *<agency, operation completed, frame-systems handled, terminal values before operation, terminal values after operation, questions in the mind, probable future operations, reasons>* tuples. The *<probable future operations>* symbolize hypotheses by **Cr**, sub-problems identified by **Re**, or suggestions by **Su**, **Se** and **Sy**.  
The *<questions in the mind, probable future operations>* parameters indicate terminals with uncertain or no slot values or incoherent granules of comprehension, and exogenously or endogenously activate specific sub-agencies, respectively. These activated agencies, run through innate algorithm trials in their *local FA* space, and then through **Log**, 'suggest' strategies towards the resolution of the *<probable future operations>* or 'suggest' new operations altogether. **Su** analyses this candidate solution space for the effective mix of partial solutions for the problem. Status updates and records of partial-solution pools in **Log** allows **Su** to backtrack and 'deliberate and reflect upon' strategies in case of erroneous or cost-ineffective choices made.
- (n) What operations are activated by the agencies depends entirely on how meanings are encoded into frames. The local *critic-selector* analyses of agency-operations, as well as global agency-suggestions are analogous to 'mentalese (Pinker 1997)' or the language of thought in the computational mind. **Log** is a manifestation of the mentalese of the computational mind.

#### 4.3.2 The working principle

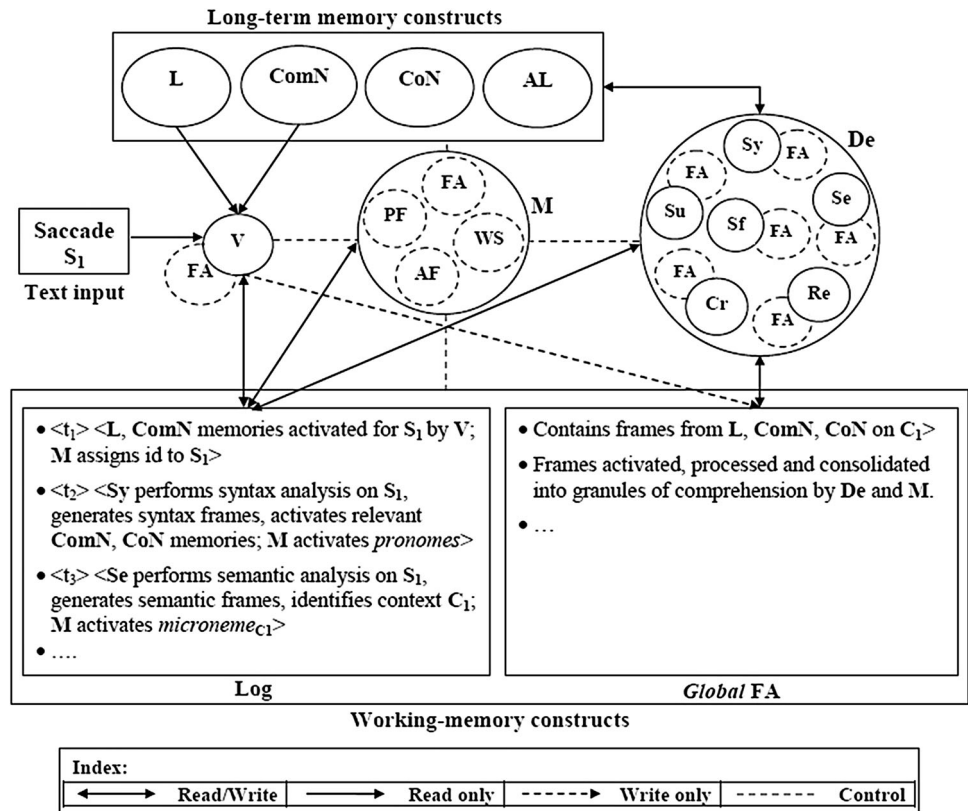
Referring to the functionalities of the defined components in the preceding section, the basic working principle of the framework (illustrated in Fig. 6), is as follows. We reckon that this principle applies to text comprehension and understanding in general as well—

Given a problem, i.e., a text to read, **V** is activated and it makes **Log** and *global FA* entries—indicating the symbols extracted, granulated and interpreted. These interpretations could include annotations like (*author\_name, text\_name, title, chapter\_name, starting words, word meanings* etc.), depending on the **L** and **ComN** memories retrieved. Once actuated, **V** extracts text in saccadic-granules (Harley 2008), the length and location of which is regulated by **De**, until reading and subsequent comprehension is complete.

All retrievals by **V** are visible, via the working-memory, to all of the other agencies to deliberate upon. The sub-agencies under **De** assess the status (familiarity, syntax, semantics, context, affects, interests, relevance, etc.) of the problem (words, clauses, sentences, paragraphs, frame-systems, etc.) and opportunistically 'suggest' interpretation mechanisms and results. These involve decomposing the problem into sub-problems and incremental-developmental iterations—through long-term to working-memory information retrievals, local frame-manipulation trials and broadcasting of predictably existing success-rendering schemes, signals to improvise upon known processes and interpretations or construct new ones from scratch, alignment of interpretations with self-interests and information consolidation—towards the formation of a granule of comprehension of the entire text sample. **M** works seamlessly in the background to support the agency activities.

Every single hypothesis, agency operation, information retrieval, changes in the working-memory is corroborated by a **Log** entry. This allows **Su** to constantly monitor (predict, visualize, question, clarify and evaluate) if the solutions provided by the different sub-agencies will eventually converge, and accordingly activate or inhibit operations (e.g. **Sy** and **Se** might be requested to re-process an incoherent granule). Ideally, an inhibited agency possesses the right to 'question' **Su**'s directions.

**Fig. 6** Illustration of the working principle of the mind-agency framework. The acronyms and connectivity lines here are to be interpreted as mentioned in Fig. 4



Thus all instructions by **Su** are annotated with encoded-reasons for evaluation and reflection. In the current version of the system, though no agency can override **Su**'s commands, none of its possible partial processing results are lost. All partially processed frames or inhibited processing vestiges can be retrieved from **PF**, on demand, for re-analysis.

An algorithmic or effective procedural view of the working principle necessitates detailed elucidation of the working-memory formats and definition of frame structures of the architecture, time and space complexity analyses, correctness and completeness verifications. As this article clearly focuses on the higher-level elements of the framework, subtle hints towards the parameters and tuples in these constructs have been provided across this article but we deliberately refrain from discussions on its fine-grained components.

Expanding on the fundamental objectives of the framework, the agency-specific functions and inter-agency activities are enumerated below. Neither is the order of the functions material (Sect. 4.3.3 elaborates on the execution modes of the architecture), nor can we be conclusive about the following being a complete list of all the cognitive functions underlying comprehension; but these do serve as a guide for the framework designers

and promote investigations on the microcosmic elements of the system.

(a) Sensory-Gateway (SG):

(a) Vision (V):

- (1) Is the visual *protospecialist* of the system, and is responsible for symbol-extraction, symbol-interpretation and symbol-granulation from saccades.
- (2) Performs morphemic analysis, i.e., the extraction of the root word, prefixes and suffixes.
- (3) References **ComN** and **L** to extract encoded meanings of morphemes; subsequent entries into **Log** activates **De**'s sub-agencies which in turn lead to retrieval of memories from **CoN** and **ComN** and **AL**.
- (4) Uses **ComN** and **L** to handle errors—prediction of incomplete text elements and ignoring stray marks.
- (5) Saccade length, and speed, time and location of retrievals are regulated by **Su** under **De**.

## (b) Deducer (De):

## (a) Syntax (Sy):

- (1) Identifies the part of speech of words, phrases or clauses or sentences using formal syntax analysis procedures, commonsense and intuition.
- (2) Creates and updates surface syntax frames, prunes inconsequential syntax frames.
- (3) Activates relevant **ComN** and **CoN** sections.

## (b) Semantic (Se):

- (1) Identifies literature type—prose or verse.
- (2) Identifies text-genre and the context from explicit or metaphorical textual cues.
- (3) Identifies the figures of speech of linguistic units.
- (4) Performs anaphora/cataphora-resolution, spatial/temporal\_sense-resolution, context-sensitive\_sense-resolution of homonyms.
- (5) Uses syntax frames to create and update semantic, thematic and narrative frames, prunes inconsequential semantic frames.
- (6) Activates relevant **ComN** and **CoN** sections.

## (c) Self (Sf):

- (1) Monitors affect progression during text processing.
- (2) Monitors belief and confidence of knowledge retrieved, or formed.
- (3) Attention or interest progression monitoring.
- (4) Identifies attachment figures of the system.
- (5) Monitors reinforcement of knowledge (over **CoN**, **ComN** and **AL**) by interaction with attachment figures or self-assessment.
- (6) Initiates upgrading of heavily reinforced **L**, **AL** and **CoN** elements to **ComN**—triggering predestined learning.
- (7) Effects recollection of memories—intuitively, ‘high-interest’ or ‘high-emotion’, or ‘high-belief’ memories

are the first ones to be retrieved from **CoN** and **ComN**.

- (8) Manipulates semantic, narrative and thematic frames.
- (9) Ensures cognitive biases do not lead to incorrect processing.
- (10) Spawns multi-mental realm reformulations of a problem; each realm in turn activates relevant agencies (**Cr**, **Re**, **Su**).
- (11) Self-reflection—judges the alignment of the text to ideals and preferences.

## (d) Recall (Re):

- (1) Retrieves memories from **ComM** and **CoN**, if all the text description parameters (e.g., *author*, *title*, etc.) extracted by **V** are known, towards emulating “automatic behavior” of the system. Else, partitions current interpretation problem into sub-problems by extrapolating with ‘similar’ experiences and context.
- (2) If all sub-problems have known solutions, activates memories of solutions in **AL** and initiates involvement of the required agencies in the text interpretation processes.
- (3) For sub-problems that have no solutions, activates **Cr**.
- (4) Activates **Su** to monitor and conquer partial solutions to an effective mechanism.
- (5) Initiates updating of **AL**, **ComN** and **ComM**.

## (e) Creativity (Cr):

- (1) Hypothesizes interpretation strategies for a given ‘new’ problem.
- (2) Evaluates differences between a problem and the ‘similar’ experiences recalled by **Re**.
- (3) Reformulates, accumulates and un-frames memories.
- (4) Transframes across contexts and memories.
- (5) Commonsense and intuitive reasoning are key reasoning tools.
- (6) Improvises upon known ‘similar’ solution strategies to counter differences—initiates solution trials by other sub-agencies.

- (7) Builds solutions from scratch—initiates transframing trials and subsequent solution trials.
  - (8) Exception handling—deals with linguistic units whose meaning cannot be ascertained from **L** or neighborhood text analysis—asks another machine, initiates web searches, asks a human, decides when to ‘give up’, etc.
  - (9) Activates **Su** to monitor solution trials to an effective mechanism.
  - (10) Initiates updating of **L**, **CoN**, **ComN** and **AL**.
  - (11) Ingenuity of solutions (cost effectiveness or new-ness) is a measure of the MIQ (Zadeh 1994), where ‘new-ness’ is relative to the system’s existing knowledge.
  - (12) Emulates ‘imagination’—the ability to visualize intentional objects (Husserl 1970).
- (f) Summary (Su):
- (1) Predicts, visualizes, questions and clarifies all computational mind activities during text processing.
  - (2) Monitors relevance and comprehension-progression through text processing.
  - (3) Generates curiosity (Gottlieb et al. 2013), questions in the computational mind, when comprehension is incomplete or unsatisfactory.
  - (4) Measures information gaps (Loewenstein 1994), attention and interest, to regulate saccade length and consequent text-intake rate by **V**.
  - (5) Instructs **V** to re-read or search for textual cues that relieve curiosity.
  - (6) Adjudges non-convergence of syntactic or semantic analyses and inhibits erroneous operations; leads to the identification of semantic errors in text.
  - (7) Consolidates solution principles of sub-problems to formulate effective text-interpretation strategies; Occam’s Razor is a notion of parsimonious problem solving, understanding and thought (Baum 2009).
  - (8) Consolidates frames resulting out of sub-problem solutions into coherent granules of facts and events.
- (9) Deliberates and reflects over successful and unsuccessful interpretations and strategies used thereof to reason or clarify success and failure.
  - (10) Reflects over inhibited processes to emulate ‘counterfactual thinking (Rose 1997)’.
  - (11) Reflections motivate ‘new’ thinking by activating **Cr** which in turn triggers other sub-agencies.
  - (12) Applies new interpretation procedures, formed by **Cr**, to problems ranked ‘similar’ by **Re**—an attempt at counterfactual thinking; motivates effectiveness tests of ‘new’ procedures against existing solutions for these problems and subsequent updating of **AL**.
  - (13) Annotates solutions with *<problem, process, result, reason>* for storage in **AL**.
  - (14) Annotates memories with *<environment descriptors, problem, solution, result, reason, affects, beliefs, etc.>* for storage in **CoN**.
  - (15) Segments text into sections—introduction, rising action, climax, resolution, and denouement, based on information, affect and interest progression.
  - (16) Restraining sub-agency operations involves backtracking through **Log** to arrive at the last ‘stable’ state of the system.
  - (17) Updates **L**, **CoN**, **ComN** and **AL**.
  - (18) Updating of **AL** triggers upgrading of the agents that symbolize algorithms under sub-agencies.
- (c) Manager (M):
- (1) The control-shell of the architecture—the hub of effective and coherent organization of agency activity; runs in the background providing housekeeping support to the inter-agency and intra-agency activities.
  - (2) System time management—System clock maintenance for **Log** entry timestamps; ensure (hard-to-soft) real-time time constraints over operations such that system cognition is at most of the order of average human cognition rates.
  - (3) Attaches unique identifiers to extracted saccadic information. These identifiers are used by **Su** to initiate verbatim recall and re-

- reading (Li et al. 2013; Payne and Reader 2006; Rothkopf 1971), and intuitively represent  $\langle \text{page\_no, location on page, keywords in neighborhood text, ...} \rangle$ .
- (4) Memory handling—long-term to working-memory placement and replacement strategies via **WS**, **AF** and **PF** to ensure thrashing avoidance and recovery, working-memory to long term memory transfers, inter long-term memory data transfers, encoding of memories (across different frame and data-types) into compressed uniform formats during storage and decoding during retrieval.
  - (5) **FA** management—maintains coherence across *local* and *global FA*, selective clearing of *local FA* (removal of only irrelevant sections), annotation of ‘trial’ and ‘applied’ results, fixed-size or adaptive (as per requirement) allotment of physical memory space for *local FA*.
  - (6) **Log**-management—read/write synchronization across multiple agencies, commit point handling (write-back all ‘correct’ short-term memory modifications to long-term memory constructs), heuristic scheduling (Erman et al. 1980) to arbitrate multiple agency-attention (**Log**-write) seeking requests.
  - (7) K-line management—to spawn or kill a K-line component (identifier-assignment, memory management, **Log** entries). *Polyneme*—tracks **FA** components denoting different ideas about a singular parent-frame (e.g. A polyneme for the parent frame ‘apple’ tracks terminals and slots for ‘color’, ‘shape’ and ‘texture’); every different sense of a homonym has a unique polyneme tracking (akin to header-nodes of linked-lists) its corresponding **FA** elements. *Micro-neme*—encodes global context parameters, as evaluated by **Se**; is used by the agencies to determine context-relevant procedures for the interpretation process. *Pronome*—handles the establishment of physical connections between frame elements, across frame-systems, across retrievals and manipulations, etc. in the **FA**. *Isonome*—simulates the same procedure across a number of things, e.g. execution of transframing procedures across multiple contexts, or the application of a ‘new’ procedure on concepts towards ‘counterfactual thinking (Roese 1997)’. *Paranome*—tracks **FA** components pertaining to an active mental realm of thinking for the given text; every active mental realm has a paranome tracking its **FA** elements.
  - (8) Context-switching (across text chapters or text sections)—involves storing the status of the current context and transferring control to a new context.
  - (9) Handle undo-redo operations dictated by **De** and consequent system state transitions—through memory, **FA**, **Log** and K-lines.
  - (10) System optimization—utilizes idle processor cycles to perform online housekeeping tasks, reflect ever system management mechanisms to reason and self-modify towards enhancement, execute **Su**’s efforts to arrive at ‘new revelations’.

#### 4.3.3 Synthesis of a computational mind

The following inferences from the agency-functionality and working principle illuminated in the preceding sections imply important synthesis issues of the framework:

- (a) The mind-agency framework is one of complex inter-agency and intra-agency connectivity; the agencies work in harmony to comprehend text or any event in the real-world.
- (b) Agency and agent construction imperatives:
  - (a) A sub-agency typically comprises of: (a) *algorithm agents* that track different methods of realizing the sub-agency functionality, (b) *function agents* that emulate typical sub-functions of the algorithm agents, and (c) *critic-selector agents* that weigh the effectiveness of different algorithms to reason and choose the best option. While the critic-selector agents monitor local appropriateness of solution strategies, **Su** monitors the global appropriateness.
  - (b) The design (Gottlieb et al. 2013) of *critic-selector agents* requires that besides monitoring the *algorithm agents*, they analyze their own competence and epistemic states, estimate their own uncertainty and execute strategies for reducing the uncertainty. This calls for understanding the physics of innate mental-rewards in the human brain that prompts information-seeking and learning towards ‘cognitive development’ in a human being and correspondingly so in an agent.
  - (c) Each agency has at least one *critic-selector agent* granule that is dedicated to the analysis of **Log** entries and subsequent agency self-activation.

- (d) The brain selects and proactively seeks out the information it wishes to sample, and this active process plays a key role in the construction of conscious perception (Gottlieb et al. 2013). Thus, global significance analysis (McCarthy 2008; Pal and Banerjee 2013)—across frames (relevance to context and comprehension) as well as interpretation strategies (co-operative and competitive effects of agency suggestions) is a crucial **Su** function.
- (e) **Re**, **Cr** and **Su** depend on the functional programming (Backus 1978) paradigm—where modules lend their functionalities towards the generation of a bespoke algorithm fitting the needs of the current text interpretation problem.
- (c) Frame handling requisites:
- (a) Each of the agencies deal with frames in one form or another.
  - (b) Solution strategies imply frame-manipulation operations; solutions imply frame-manipulation results.
  - (c) Frame manipulation necessitates the definition of a calculus or a frame manipulation language, leading to the formation of conceptual associations.
  - (d) Frame manipulation schemes need to seamlessly integrate and operate across multiple data-types representing different sensory memories.
  - (e) Besides parameters that describe a fact or an event, frames need to embody parameters that define the system's belief of the world and itself. The Z-number (Banerjee and Pal 2013; Pal et al. 2013; Zadeh 2011) philosophy is an effective strategy towards the representation of subjective beliefs.
  - (f) Any concept has two simultaneous representations—integrated (after frame-transframing and frame-unframing operations) and differentiated (after frame-accumulating operations) [Refer to Sect. 3.2.1 for frame operations].
  - (g) Typical frame states are:
    - (1) *Activation* Recall of frames, terminals and slot values associated with the current text stimulus. On activation, terminal slots are filled with 'default (intuitive)' or 'most likely [high-certainty (Banerjee and Pal. 2013; Pal et al. 2013)]' values for the terminal.
    - (2) *Instantiation* Assignment of slot values particular to the current stimulus; an activated terminal is instantiated if the existing slot value is updated to reflect the current text.
- (d) A rule of thumb for the time frame for **WS** is roughly of the order of the time for processing a paragraph, while that for **AF** is of the order of time for processing a page. **M** tracks the approximate time to process an average paragraph or page and modulates the time window accordingly.
- (e) During reading, the human brain typically (McCarthy 2008):
- (a) Processes words in the text in the 'foreground'.
  - (b) Unconsciously takes into account the ambient lighting, the seating comfort, the time, the arrival of people, ambient sounds, i.e., the brain processes these elements in the 'background', and these environmental descriptors can often ['incorrectly (Ariely 2008; Banaji and Greenwald 2013)'] influence the interpretation of the text.
  - (c) The 'foreground' and the 'background' processing activities work in tandem and take place when the reader is actually reading (online) or mulling over the read text (offline). Thus, while in this article we have restricted to just a description of **V**, each of the sub-agencies under **SG** of a computational mind plays a critical role in text understanding. The important difference that our system has with the human mind is that **Sf** has been delegated an essential task of immunizing interpretations from cognitive biases; thus **Sf** tries to balance between emotional and rational thinking.
  - (f) Drawing from point (e), a computational mind ideally operates in the following modes (Fig. 7, presents a snapshot of the operation modes of a computational mind, alluding to Minsky's divisions of the layers of thinking (Minsky 2006), and Fig. 8, elaborates on the same):
    - (a) Based on principles of dynamicity (current  $\text{time\_frame} = t$ ):
      - (1) *Online processing* ( $t$ ) (Seth et al. 2006) Processing stimulus that is active at  $t$ ; is analogous to the 'experiencing self (Kahneman 2011)'. This mode represents conscious association formation due to transactions between organisms and environments.
      - (2) *Offline processing* ( $t$ ) (Seth et al. 2006) Processing stimulus that was active at

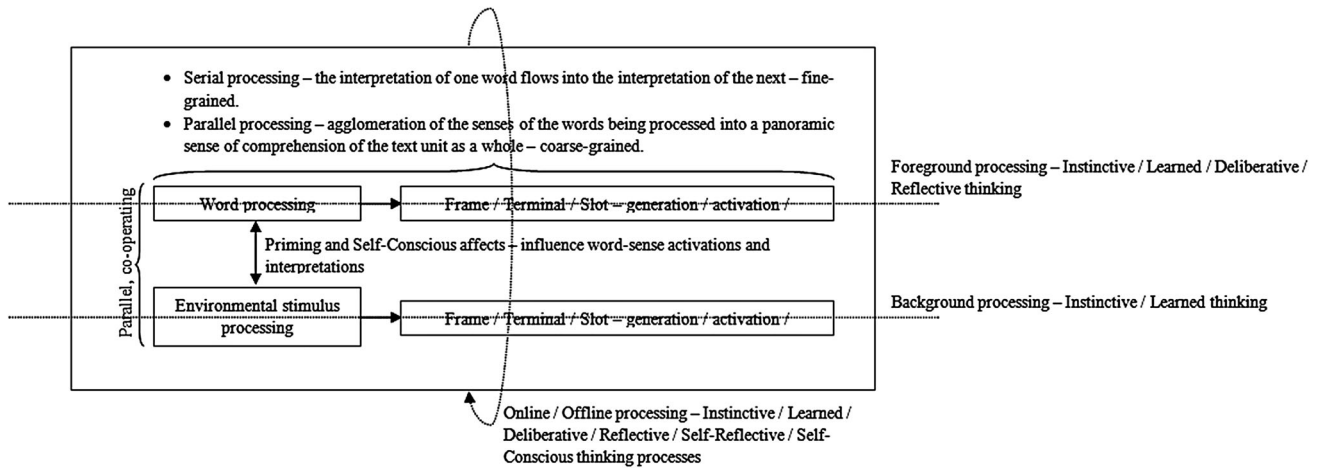


Fig. 7 A snapshot of the operation modes of a computational mind during text processing

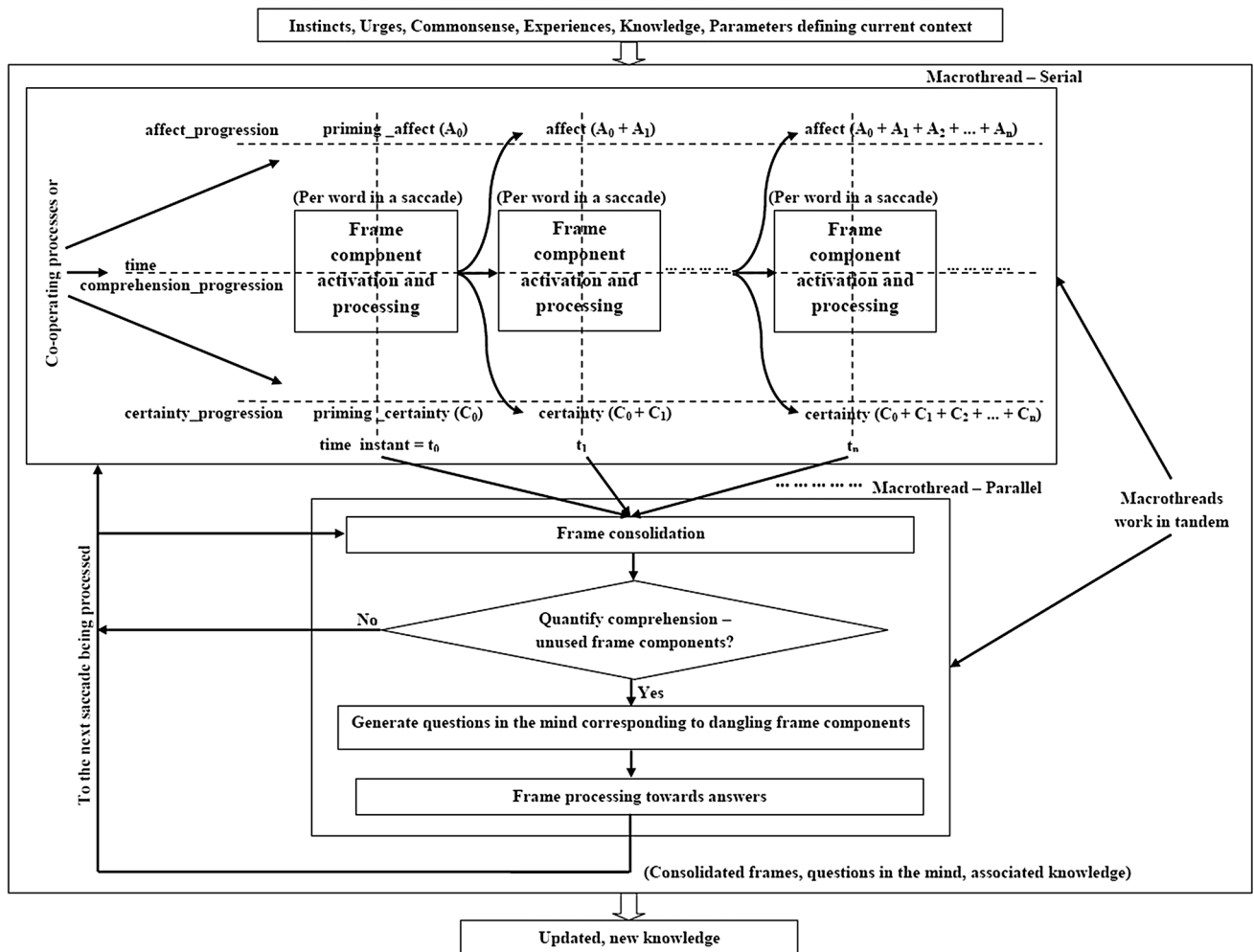


Fig. 8 A detailed illustration of the operation modes in sync with frame-processing for text understanding in a computational mind

some previous time frame ( $<t$ ), but is no longer active at  $t$ . This mode represents the action of 'mulling over' or 'reflection', and is analogous to the 'remembering self (Kahneman 2011)'. It represents conscious association formation during dreaming, reverie, abstract thought, planning, or imagery.

- (b) Based on differences in conscious-processing activities of stimuli at time ( $t$ ):
- (1) *Foreground processing ( $t$ )* Activation or instantiation of frame units by actual or the intended stimuli ( $S$ ) at  $t$ ; symbolizes conscious mind activity.
  - (2) *Background processing ( $t$ )* Activation or instantiation of frame units by environmental or commonsense cues while processing  $S$  at  $t$ ; symbolizes sub-conscious or unconscious mind activity.
- (c) The above modes can be further categorized into:
- (1) *Serial processing* Where the outcome of processing an element at time  $t$  flows into the processing of an element at time ( $t + 1$ ) or later. For example, the outcomes of online processing activities serve as inputs during the offline processing phase—on a global scale, or the interpretation of a saccade of text effects the interpretation of the succeeding ones—on a local scale. Conscious activities are serial (Baars 1988).
  - (2) *Parallel or co-operative processing* Where a number of stimuli are processed in tandem. For example, the foreground and the background processing phases work in sync towards the comprehension of the present context (von Neumann 2012). Unconscious activities are parallel (Baars 1988).

A computational mind ideally, not only performs serial and parallel processing simultaneously, but the outcomes of these processing activities co-operate with each other as well—acknowledging the simultaneous left and the right brain processes across the corpus callosum. For example, while the eyes

serially extract saccades of text, the words in each saccade are concurrently co-operatively processed and the interpretation of one saccade flows into the next—leading to an incrementally growing module of comprehension. We refer to these modes as **macrothreads**.

Each active serial or parallel macro-thread is further composed of a number of **microthreads**. Considering reading, the microthreads involved in a serial macrothread are the extracted saccades and environmental inputs at time ( $t$ ), while those for the parallel macrothreads are the individual words in a saccade or multiple active saccades. These operation modes are in line with the concept of 'thinking without thinking (Gladwell 2005)'.

- (g) A primitive granule processed by the human vision (Cristobal et al. 2011) system is the text contained in a saccade (Harley 2008). Following experimental studies in Miller (1955), it perhaps is right to conclude that a saccade has a maximum of seven words. Now if a machine were to process a seven-word saccade, it should be able to activate seven threads for concurrent co-operative handling of the intra-saccade microthreads, as well as additional threads for handling concurrent co-operative processing of the inter-saccade macrothreads. Considering typical present day processor architectures, a saccade with more than seven words could perhaps be easily accommodated. This data-driven design perspective reflects a conscious shift away from the 'word-at-a time (Backus 1978)-thinking philosophy' underlying von-Neumann bottlenecks.

## 5 Analysis of the framework

Having described the components that constitute the mind-agency framework, this section focuses on analyzing its correctness and completeness—in terms of the theories it is based upon. These evaluations here are only in terms of us having identified all the requisite functions and modules. The design shall only be complete once the agents, data structures and knowledge bases are in place, whereupon the agencies are functional and execute as per design expectations.

We present here a dry-run through the working principle the outputs of which have been validated by human subjects, followed by a study of correspondences with

structures in the human brain and layers in Minsky's model of the human mind. The framework is then conceptually compared with existing cognitive frameworks.

### 5.1 A dry-run test of the framework

Table 3 presents an explicit run through the framework—depicting the stages of comprehension and the roles of the mind-agencies. Components in the table abide by the following schematics:

- (a) <bold> indicates 'frame header'.
- (b) <italics> indicates 'terminal'.
- (c) <bold and italics> indicates 'slot value'.
- (d) () indicates 'frame-terminal' connectivity relation.
- (e) Arrow heads indicate connectivity destinations; destinations could be 'terminals' or 'slot values'.

*Assumptions* Each of the mind-agencies and memory constructs is functional, as per the descriptions in Sect. 4.  
*Input text* A duck waddled past the post-box. It didn't notice a cat nearby.

*Expected output* A narrative of comprehension—summarizing the surface and deep semantics of the input text.

#### Observations

- (1) Inference results and data of one stage percolate down to the next stage of comprehension.
- (2) Entries across time units indicate **Log** as well as *global FA* values.
- (3) Each time\_unit-action\_thread intersection implies a macrothread, while the entries within the intersection symbolize constituent microthread operations.
- (4) Activities of **M** have not been deliberately highlighted, as we wanted to focus exclusively on the phenomenon of comprehension.
- (5) At time  $T_5$ , **Su** summarizes the surface semantics (depicted in the darkened table entry) of the text input. All that follows are results of thinking across the four higher layers of the mind.
- (6) The progression of comprehension depicted above was validated by the thought processes of fifteen random individuals. These individuals were asked to list—over a time period of 2 days—all that their minds processed in relation to the given text input, and in the order that their thoughts were activated. Results of surface semantics and reflective assumptions matched with twelve of the test subjects; the remaining, unfortunately did not process beyond surface semantics.

We are currently in the process of performing more such experiments, so as to understand better the

average thought processes given random text instances.

- (7) **Sf** is biased by the 'availability' heuristic, at time  $T_5$  where it assumes a pessimistic perspective over <predator-prey>. **Cr** at Time  $T_{12}$  suggests an optimistic viewpoint.
- (8) The results above highlight those due to online-foreground processing. Offline or background processing could include views on the <post-box>; **Cr** prescription strategies like <duck flying onto post-box> as an <escape> mechanism and so on.
- (9) From the above example, it is evident that our framework is conceptually a cognitive model of text comprehension, as it demonstrates: (a) multiple-realm 'thinking', (b) ambiguity resolution, (c) recollection and reflection, and (d) subjective decision-making.
- (10) The question of importance at this juncture is when will it be evident that a machine was 'thinking' or behaving 'intelligently'? Drawing from Ryle (1949), the procedures that the machine uses in order to arrive at solutions is an indication of its intelligence, where the procedures are an amalgamation of its knowledge, intuition, commonsense, and experience. Furthermore, considering the implication of the self-consciousness method of thinking, (Seth 2010; Seth et al. 2006) indicate the need for effective means for the measurement of conscious thoughts and states by a machine. These methods need to incorporate both objective and subjective machine responses. The interesting question here is would a 'conscious' 'thinking' 'understanding' machine be immune to consciousness disorders leading to psychiatric or neurologic disorders or minimal conscious states?
- (11) Referring to (Erman et al. 1980) for the key requirements of knowledge based language understander systems:
  - (a) Representation and structuring of the problem in a way that permits decomposition.
  - (b) Total interpretation is to be broken down into hypotheses and modularized into different types of knowledge that can operate independently and co-operatively.

The following features of the framework support the conceptual acknowledgement of these requirements:

- (a) Not only have we factored text comprehension into its component functions (Sect. 4.2) and assigned their execution to mind-

**Table 3** The action dynamics of comprehension by a computational mind

Time	Action threads
T <sub>0</sub>	V—Extracts: duck, waddles, past, post-box, ‘;’, it, didn’t notice, cat, nearby, ‘;’
T <sub>1</sub>	V—Granulates <duck, waddles, past, post-box, ‘;’>, <it, didn’t, notice, cat, nearby, ‘;’>
T <sub>2</sub>	S <sub>y</sub> —Activates syntax frames for S <sub>1</sub> <duck> → <noun> <verb> <adjective> <waddles> → <verb> <past> → <adjective> <noun> <preposition> <adverb> <post-box> → <noun>
T <sub>3</sub>	S <sub>y</sub> —Activates syntax frames for S <sub>2</sub> <it> → <pronoun> <didn’t> → <contraction> → <did not> <did not> → <negative auxiliary verb> <notice> → <verb> <noun> <cat> → <noun> <verb> <nearby> → <adjective> <adverb>
T <sub>4</sub>	S <sub>u</sub> consolidates frames for S <sub>1</sub> into a narrative <duck complementarily crossing a post-box>
T <sub>5</sub>	S <sub>u</sub> consolidated frames for S <sub>1</sub> and S <sub>2</sub> into a narrative (<duck complementarily crossing post-box> && <there is cat near post-box> && <duck did not see cat>)

M—Assigns granules id S <sub>1</sub> and S <sub>2</sub> , respectively	V—References L and ComN and retrieves knowledge on: <duck>, <waddle>, <past>, <post-box>, <didn’t>, <notice>, <cat>, <nearby>
Re recognizes from S <sub>1</sub> : <duck> → (motion) → <waddle>	Sf assigns for S <sub>1</sub> : <affect> → <complacent>
Se—Prunes irrelevant syntax frames for S <sub>1</sub> <duck> → <noun> <waddles> → <verb> <past> → <preposition> <post-box> → <noun>	Se—Constructs semantic frames for S <sub>1</sub> Frame1 <sub>S1</sub> —<duck> ← (what) ← <waddled> → (where) → <past>; Frame2 <sub>S1</sub> —<past> → (what) → <post-box>
Se—Prunes irrelevant syntax frames for S <sub>2</sub> <it> → <pronoun> <did not> → <negative auxiliary verb> <notice> → <verb> <cat> → <noun> <nearby> → <adverb>	Se—Constructs semantic frames for S <sub>2</sub> Frame1 <sub>S2</sub> —<it> → (whom) → <duck>; Frame2 <sub>S2</sub> —<duck> ← (who) ← <did not> → (what) → <notice>; Frame3 <sub>S2</sub> —<notice> → (what) → <cat> Frame4 <sub>S2</sub> —<cat> → (where) → <nearby> Frame5 <sub>S2</sub> —<nearby> → (reference point) → <postbox>
Re recognizes from S <sub>2</sub> <duck> → (what) → <bird> <cat> ← (what) ← <predator> → (of) → <birds>	Sf analyzes for S <sub>2</sub> <predator—presence> → (affect on prey) → <tense> <tense> → (natural reflex of prey) → <escape trial>
	Sf adjudges status “highly interesting” and focuses attention on the <predator—prey> subject

Table 3 continued

Time	Action threads
T <sub>6</sub>	<b>Su</b> stimulates <b>V</b> to extract new saccades
T <sub>7</sub>	<b>V</b> reports no new text to read
T <sub>8</sub>	<b>Su</b> reflects and arouses curiosity If the cat noticed the duck, will the cat catch the duck? Can the duck escape? Will the duck ever notice the cat? What if they do not notice each other? <b>Sf</b> reflects and activates realms <b>R<sub>1</sub></b> —Duck notices cat first <b>R<sub>2</sub></b> —Cat notices duck first <b>R<sub>3</sub></b> —Cat and duck notice each other at the same time <b>R<sub>4</sub></b> —Neither notice each other
T <sub>9</sub>	<b>Re</b> recalls for <b>R<sub>1</sub>, R<sub>3</sub></b> < <b>bird</b> > → (activity) → < <b>fly</b> > < <b>run</b> > < <b>swim</b> > < <b>jump</b> > <...> < <b>duck</b> > → (can) → < <b>fly</b> > < <b>cat</b> > → (cannot) → < <b>fly</b> > < <b>fly</b> > → (action result) → < <b>escape</b> >
T <sub>10</sub>	<b>Re</b> recalls for <b>R<sub>2</sub>, R<sub>3</sub></b> < <b>cat</b> > → (can) → < <b>pounce</b> > < <b>duck</b> > → (can) → < <b>jump</b> > < <b>cat pounce</b> > ← (of) ← < <b>greater height</b> > < <b>duck jump</b> > → (than) → < <b>duck jump</b> > < <b>duck jump</b> > → (action result) → < <b>not escape</b> >
T <sub>11</sub>	<b>Re</b> recalls for <b>R<sub>1</sub>, R<sub>3</sub></b> < <b>duck</b> > → (can) → < <b>run</b> > < <b>duck</b> > → (can) → < <b>run</b> > < <b>cat run</b> > ← (of) ← < <b>faster speed</b> > < <b>duck run</b> > → (than) → < <b>duck run</b> > < <b>duck run</b> > → (action result) → < <b>not escape</b> > <b>Re</b> recalls for <b>R<sub>1</sub>, R<sub>3</sub></b> < <b>duck</b> > → (can) → < <b>swim</b> > <b>Su</b> inhibits <b>Re</b> , reason < <b>swim</b> > → (place) → < <b>water body</b> > < <b>water body</b> > → < <b>pond</b> > < <b>lake</b> > < <b>sea</b> > <...> No < <b>water</b> > frame active <b>Su</b> arouses curiosity Is there water body nearby?
T <sub>12</sub>	For <b>R<sub>1</sub></b> and <b>R<sub>3</sub></b> , <b>Su</b> summarizes (< <b>duck notices cat</b> > && < <b>duck flies away</b> >) For <b>R<sub>2</sub>, Su</b> summarizes (< <b>cat notices duck</b> > && < <b>duck is hurt</b> >) <b>Su</b> activates <b>Cr</b> for <b>R<sub>2</sub></b> , reason Summary against <b>Sf</b> preference <b>Cr</b> suggests < <b>Optimism</b> > → < <b>cat does not hurt duck</b> >

agencies (Sect. 4.3.1), these agencies are further composed of agents that decompose these functions into algorithmic steps (Sect. 4.3.3).

- (b) Deconstruction of interpretations into hypotheses and knowledge modularization is supported through–
  - (1) **Re** decomposes an interpretation problem into “similar” sub-problems and recalls known solutions.
  - (2) **Cr** hypothesizes new solution perspectives.
  - (3) **Su** periodically summarizes comprehension statuses which in turn activates solution suggestions by different agencies.
  - (4) *Critic-selector agents* critically analyze multiple approaches towards the realization of an agency-function.
  - (5) **Global FA** and **Log** serve as global workspaces for the agencies to cooperate towards solutions.
  - (6) **Local FA** supports independent agency-activity trials, moderated by *critic-selector agents*.
  - (7) As is evident from the hierarchy of the **De** agencies (Sect. 4.3.1), these operate across a number of information-granular levels.
  - (8) Multiple solutions across agencies form pools of candidate partial solutions, which **Su** combines into effective global solutions.
  - (9) Knowledge—modularized into facts, concepts, intuition, commonsense and procedures are referenced by agencies, relative to the demands of the status of comprehension.

## 5.2 Correspondence between the mind-agencies and brain-functions

Table 4 summarizes the analogy between the brain lobes in the cerebral cortex and the mind-agencies, and Table 5 depicts the one-to-one correspondence between the memory categories of the human brain and the memory constructs of the framework. By virtue of the total coverage of the lobes and the memories by the mind-agencies and memory structures, respectively, we consider our design complete.

## 5.3 Correspondence between the mind-agencies and the layers of the mind

The functions of the agencies are indicators of the layers of the human mind that they embody, and we summarize the correspondence in Table 6. Evidently, the function boundaries are not crisp and each agency covers more than one layer of the human mind. By the strength of the total coverage of the functionalities across the layers by the mind-agencies, we consider our design complete.

## 5.4 Comparison with Hearsay and ‘conscious’ software agents (CMATTIE, IDA)

This segment briefly elucidates on the conceptual similarities and differences with existing ‘intelligent’ ‘reflective’ ‘conscious’ agents—Hearsay (Erman et al. 1980), CMATTIE (McCauley et al. 2000; Zhang 1998) and IDA (Baars 1988; Franklin 2003). Our framework draws from the many advantages of these systems (described in Sect. 2) and aims to augment their abilities towards a truly intelligent machine (Turing 1950).

### Similarities–

Each of the existing agents and our mind-agency architecture:

- (a) Are based on the ‘Society of Mind’ theory.

**Table 4** Correspondence between the cerebral cortex regions and the mind-agencies, based on their functional analogy

Cerebral cortex region	Framework agencies and functions
Occipital	<b>V</b>
Frontal	Broca’s area— <b>Sy</b> Self definition, attention, social behavior— <b>Sf</b> Reasoning, judgment, strategic thinking— <b>Re, Cr, Su</b>
Parietal	Angular gyrus— <b>Se, Sf, Su, Re, Cr</b>
Temporal	Wernicke’s area— <b>Se</b> Amygdala— <b>Sf, Su</b> Hippocampus— <b>Su</b> Basal Ganglia— <b>Sf, Su, Re</b> Recognition— <b>Re</b>

**Table 5** Correspondence between categories of the human memory and the memory constructs of the framework

Human memory	Framework memory constructs
Working	<i>global FA</i> ; <i>local FA</i> of <b>De</b> and <b>M</b> sub-agencies; <b>AF</b> ; <b>PF</b> ( <b>WS</b> $\subseteq$ <b>AF</b> and is therefore not explicitly mentioned)
Declarative	<b>CoN</b> ; <b>AL</b>
Procedural	<b>ComN</b>
Long-term	<b>CoN</b> ; <b>ComN</b> ; <b>AL</b>
Short-term	First set of entries into <i>global FA</i> by <b>SG</b>
Sensory	<i>local FA</i> of <b>SG</b> sub-agencies
Visual, Olfactory, Haptic, Taste, Auditory	Memories annotated by the senses they pertain to—indicated by their data-types in <b>ComN</b> and <b>CoN</b>
Autobiographic	Subset of <b>CoN</b>
Retrospective, Prospective	To be constructed out of <b>ComN</b> and <b>CoN</b> (Intuitively, <b>PF</b> could be instrumental in the emulation of these memories)

**Table 6** The participation of the agencies in the thinking process

Layers of thinking In the human mind	Computational mind-agencies							
	V	Sy	Se	Sf	Re	Cr	Su	M
Instinctive reactions Accept text-input through the appropriate sensory organs	*							*
Learned reactions Assign meaning to the elements seen—alphabets, digits, special symbols, white-spaces, punctuation; agglomeration of symbols into words, numbers, codes, phrases, clauses, sentences; syntax and semantic analysis of the text extracted; literature categorization into prose, poem, etc.; genre resolution	*	*	*		*		*	*
Deliberative thinking Disambiguation of word-meanings, sentence-meanings, genres; rhetoric and prosodic analysis; analyze relevance and coherence of flow of concepts across text; consolidate individual text-elements into concepts; visualize scenes		*	*		*	*	*	*
Reflective thinking Reason and optimize deliberative thinking processes; generate curiosity (questions in the computational mind) and activate schemes to gratify the same; build cross-text and cross-contextual associations		*	*		*	*	*	*
Self-reflective thinking Evaluate interest and comprehension progression through text; overcome cognitive biases and reform concepts; text section identification—introduction, rising action, climax, denouement and conclusion; regulate eye-tracking (re-read sections, reading speed)				*			*	*
Self-conscious emotion Attachment of emotions or levels of interest and perceptions to the entire text; to what extent does the text come-up to the reader's expectations and ideals—is it taboo, inspirational, fun, tragic, unputdownable, etc.; will the reader recommend it to anyone; will the reader read it again; how does the current reading affect the reader—did the reader gain new knowledge, which concepts were clarified				*			*	*

An asterisk indicates participation; the agency-name codes here are the same as that for Fig. 4]

- (b) Are deliberative and reflective—over mechanisms that realize the purpose of their design (e.g. speech understanding).
  - (c) Rely on co-operative concurrent processing activities across modules for voluntary action selection and constraint satisfaction.
  - (d) Exhibit learning.
  - (e) Exhibit affects.
  - (f) Do not represent 'forgetfulness' or mechanisms to handle internal or external distractions.
- Differences, or rather, distinctive conceptual enhancements in the proposed mind-agency framework—

Elements present in our mind-agency architecture but absent in the existing architectures:

- (a) Acknowledgement of ‘automatic’ or intuitive behavior—based on the principle of continued reinforced learned behavior towards ‘automatic’ impulses or thinking without thinking (Gladwell 2005). This should predictably prevent the entire complex framework being activated for trivial language units (a key disadvantage of the existing systems).
- (b) Acknowledgement of commonsense reasoning.
- (c) The notion of the machine ‘self’—‘self-reflection’ and ‘self-consciousness’—towards machine ‘neurogenesis (Chugani et al. 2001)’ and subjective decisions (regulation of **V**, condition comprehension, etc.)
- (d) The concept of encoding ‘reasons’ for failure and success of solution or interpretation strategies—given a context and the section of text being processed, thereby laying the foundations for possible ‘self-modification’ or ‘self-evolution’ across essential system functions (those undertaken by **M**) towards system optimization.

Having identified the agencies and their corresponding operations, the next stage of the modeling process calls for the following tasks, and working towards these are where our future intentions lie:

- (a) Formalization of the framework—describing its functions in simple computational terms (Backus 1978).
- (b) Identification and enumeration of the agents and their functions under each of the agencies.
- (c) Specifics of all the memory constructs, frame-formats and frame-manipulation strategies.
- (d) Our design is clearly hybrid—incorporating both symbolic and connectionist features. A deeper insight into this is needed.
- (e) The ultimate challenge of our model lies in testing its robustness in dealing with the dogmas of language comprehension (Clark 1997).
- (f) Identification and formalization of parameters that define the ‘self’.

## 6 Conclusion

Right through antiquity down to the twenty-first century, thinkers, philosophers and scientists have spent years in trying to solve the ‘mysteries’ of the human brain—‘What is the ‘mind’ and how or why does it act the way it does?’ ‘How does the mind lead to intelligence?’ ‘What is it that differentiates a ‘normal’, an ‘afflicted’ and a ‘genius’ mind?’ ...

With the advent of research streams pertaining to linguistics, neuroscience, artificial intelligence, psychology,

and cognition, answers to which parts of the brain are activated in response to specific stimuli and abstract concepts on how the mind functions have been unearthed. But an accurate, scientific definition of the ‘mind’ and its functions remains elusive.

The investigations illustrated here do not, in any way, reveal answers to the questions mentioned above, but does attempt at contributing to the ‘thinking-machines’ research initiatives heralded by Turing (1950). What Turing pioneered, through this phenomenal article, is the need to think about ‘thinking’ in a disciplined way and view the mind as a scientific phenomenon involving countably infinite moving parts—visualizing the mind as a society of interacting agents.

This article is a treatise on our first steps towards the realization of a novel ‘cognitive’ model of text comprehension, based on the ‘Society of Mind (Minsky 1986)’ and the ‘Emotion Machine (Minsky 2006)’ theories, and key elements of existing language understanders. Not only does our model look into emulating the key steps in reading and comprehension, like eye-tracking etc., but also aims at incorporating the concept of ‘thinking’ across multiple realms towards arriving at text-visualizations. We describe here the top-level components of the architectures, without divulging their fine-grained technicalities, followed by a discussion on its working theory and realization constraints.

Major discoveries and hard work lie ahead before we uncover a foundation for a computational mind that is anything as basic like the chromosomes, genes and genetic code. We do not claim that our proposed design mimics the vast repertoire of mind functions nor have we defined every psychological process in its computational equivalents, but we have here a set of very basic agencies that work in unison and harmony to realize textual understanding. The concepts here serve as a blueprint for our continued evolutionary design initiatives. What we envision is that the design, instead of imitating any of the authors or people we know, be able to define its own self, be self-organized, dynamic, adaptable, and social—the mark of a truly intelligent object, as defined in McCarthy (1995, 2008).

A cognitive model of text understanding, we believe, applies to the development of ‘intelligent’ and ‘symbiotic’ man-machine interactive systems—capable of ‘understanding’ deep semantics—plagiarism-checkers, library cataloguing systems, text summarizers, differential diagnosis systems, educational aids for children with reading disorders, etc. Extending the model to include comprehension of language in all its forms is our ultimate goal.

Interestingly, this project presents an opportunity for introspection on ‘self’ and acknowledgement of oneself as a ‘thinker’ towards understanding the innate ‘algorithms’ guiding daily activities. Thus, besides the engineering

perspective, the research involved herein, has profound philosophic ramifications as well.

We shall not cease from exploration and the end of all our exploring will be to come back to the place from which we came and know it for the first time—T.S. Eliot.

**Acknowledgments** This project is being carried out under the guidance of Professor Sankar K. Pal who is an INAE Chair Professor and J.C. Bose Fellow of the Government of India. The authors acknowledge Alan Turing as the prime inspiration for the work described herein.

## References

- Ariely D (2008) Predictably irrational: the hidden forces that shape our decisions. Harper Collins, NY
- Ashby WR (1952) Design for a brain. Butler and Tanner Ltd., London
- Baars BJ (1988) A cognitive theory of consciousness. Cambridge University Press, Cambridge
- Baars BJ (1997) In the theater of consciousness: the workspace of the mind. Oxford University Press, Oxford
- Baars BJ (2002) The conscious access hypothesis: origins and recent evidence. *Trends Cogn Sci* 6(1):47–52
- Backus J (1978) Can programming be liberated from the von Neumann style? A functional style and its algebra of programs (ACM Turing Award lecture). *Commun ACM* 21(8):613–641
- Baddeley AD (1966) The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quart J Exp Psychol* 18(4):302–309
- Banaji MR, Greenwald AG (2013) Blindspot: hidden biases of good people. Delacorte Press, NY
- Banerjee R, Pal SK (2013) The Z-number enigma: a study through an experiment. In: Yager RR, Abbasov AM, Reformat MR, Shahbazova SN (eds) *Soft computing: state of the art theory and novel applications*, vol. 291 of studies in fuzziness and soft computing, Springer, Berlin/Heidelberg, pp 71–88
- Baum EB (2009) Project to build programs that understand. In: Goertzel B, Hitzler P, Hutter M (eds) *In: Proceedings of second conference on artificial general intelligence*, vol. 8 of advances in intelligent systems research, Atlantis Press, Paris, pp 1–6
- Bobrow DG (1964) Natural language input for a computer problem solving system. PhD thesis, Massachusetts Institute of Technology
- Brains in Silicon. <http://www.stanford.edu/group/brainsinsilicon/index.html>. Accessed 8 April 2014
- Bush V (1945) As we may think. *Atl Mon* 176(1):101–108
- Charniak E (1972) Toward a model of children's story comprehension. Technical report, MIT Artificial Intelligence Laboratory
- Chomsky N (1959) A review of B.F. Skinner's "verbal behavior". *Language* 35(1):26–58
- Chomsky N (1991) Linguistics and cognitive science: problems and mysteries. In: *The chomskyan turn*, Blackwell Publishing, Oxford, pp 26–53
- Chugani HT, Behen ME, Muzik O, Juhász C, Nagy F, Chugani DC (2001) Local brain functional activity following early deprivation: a study of post institutionalized Romanian orphans. *NeuroImage* 14(6):1290–1301
- Clark HH (1997) Dogmas of understanding. *Discourse Process* 23:567–598
- Conway MA, Pleydell-Pearce CW (2000) The construction of autobiographical memories in the self-memory system. *Psychol Rev* 107(2):261–288
- Cristobal G, Schelkens P, Thienpont H (eds) (2011) *Optical and digital image processing: fundamentals and applications*. Wiley-VCH Verlag GmbH and Co., KGaA, Weinheim
- Dennett DC (2013) The normal well-tempered mind. <http://www.edge.org/conversation/the-normal-well-tempered-mind>
- Eccles JC, Ito M, Szentagothai J (1967) *The cerebellum as a neuronal machine*. Springer-Verlag, NY
- Erman LD, Hayes-Roth F, Lesser VR, Reddy DR (1980) The Hearsay-II speech-understanding system: integrating knowledge to resolve uncertainty. *ACM Comput Surv* 12(2):213–253
- Ferrucci D, Brown E, Chu-Carroll J, Fan J, Gondek D, Kalyanpur AA, Lally A, Murdock JW, Nyberg E, Prager J, Schlaefer N, Welty C (2010) Building Watson: an overview of the DeepQA project. *AI Mag* 31(3):59–78
- Franklin S (2003) IDA: a conscious artifact? *J Conscious Stud* 10:47–66
- Franklin S, Patterson FG (2006) *The LIDA architecture: adding new modes of learning to an intelligent, autonomous, software agent*. Integrated Design and Process Technology, San Diego
- Gladwell M (2005) *Blink: the power of thinking without thinking*. Little Brown and Company (Hachette Book Group), NY
- Gottlieb J, Oudeyer P, Lopes M, Baranes A (2013) Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends Cogn Sci* 17(11):585–593
- Grosz BJ (2012) What question would Turing pose today? *AI Mag* 33(4):73–81
- Harley TA (2008) *The Psychology of Language: From Data to Theory*, 3rd edn. Psychology Press—Taylor and Francis Group, New York
- Harrison H, Minsky M (1992) Unpublished chapters of "The Turing Option". <http://web.media.mit.edu/~minsky/papers/option.chapters.txt>
- Havasi C, Speer R, Alonso J (2007) Conceptnet 3: a flexible, multilingual semantic network for common sense knowledge. In: *Proceedings of recent advances in natural language processing*, pp 27–29
- Hayes-Roth B (1985) A blackboard architecture for control. *Artif Intell* 26:251–321
- Hewitt C (1970) *Planner: a language for manipulating models and proving theorems in a robot*, Massachusetts Institute of Technology—Project MAC—Artificial Intelligence—Memo 168, August 1970
- Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev* 80(3):953–978
- Hunt J (2002) *Blackboard architectures*. Technical Report 1, JayDee Technology Ltd., Wiltshire
- Husserl E (1970) *Logical investigations* (Translated from German). Routledge and Kegan Paul Ltd, London
- Jankowski A, Skowron A, Swiniarski RW (2013) Interactive complex granules. In: Szczuka MS, Czaja L, Kacprzak M (eds) *CS & P*, vol. 1032 of CEUR workshop proceedings, vol 1032, pp 206–218. [CEUR-WS.org](http://www.ceur-ws.org)
- Kahneman D (2011) *Thinking, fast and slow*. Farrar, Straus and Giroux, NY
- Kofka K (1935) *Principles of gestalt psychology*. Lund Humphries, London
- Kokinov BN (1994) The dual cognitive architecture: a hybrid multi-agent approach. In: Conn A (ed) *Proceedings of 11th european conference on artificial intelligence (ECAI)*, John Wiley and Sons, Ltd, pp 203–207
- Kokinov B (1989) About modelling some aspects of human memory. In: *Man-computer interaction research (MACINTER-II)*, Elsevier, Amsterdam, pp 349–359
- Kowalski R (2011) *Computational logic and human thinking: how to be artificially intelligent*. Cambridge University Press, NY

- Langley P, Laird JE, Rogers S (2009) Cognitive architectures: research issues and challenges. *Cogn Syst Res* 10(2):141–160
- Li L, Chen G, Yang S (2013) Construction of cognitive maps to improve e-book reading and navigation. *Comput Educ* 60(1):32–39
- Lieberman H, Liu H, Singh P, Barry B (2004) Beating common sense into interactive applications. *AI Mag* 25(4):63–76
- Lin TY (1997) Granular computing. Technical report, Announcement of the BISC special interest group on granular computing
- Liu H (2004) Montylingua: an end-to-end natural language processor with common sense. [web.media.mit.edu/~hugo/montylingua](http://web.media.mit.edu/~hugo/montylingua)
- Loewenstein G (1994) The psychology of curiosity: a review and reinterpretation. *Psychol Bull* 116(1):75–98
- Maes P (1987) Concepts and experiments in computational reflection. In: Meyrowitz NK (ed) Proceedings of conference on object-oriented programming systems, languages and applications (OOPSLA), ACM, NY, pp 147–155
- Majumdar A, Sowa J, Stewart J (2008) Pursuing the goal of language understanding. In: Eklund P, Haemmerlé O (eds) Proceedings of 16th international conference on conceptual structures: knowledge visualization and reasoning, Springer-Verlag, Berlin, pp 21–42
- McCarthy J (2008) The well-designed child. *Artif Intell* 172(18):2003–2014
- McCarthy J (1995) Making robots conscious of their mental states. In: Machine intelligence, Oxford University Press, NY, pp 3–17
- McCarthy J (1959) Programs with commonsense. In: Semantic information processing, MIT Press, MA, pp 403–418
- McCauley L, Franklin S, Bogner M (2000) An emotion-based “conscious” software agent architecture. In: Paiva A (ed) Affective interactions, vol. 1814 of lecture notes on artificial intelligence, Springer-Verlag, Berlin, pp 107–120
- McGaugh JL (2004) The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu Rev Neurosci* 27:1–28
- Mead C (1990) Neuromorphic electronic systems. *Proc IEEE* 78:1629–1636
- Miller GA (1955) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 101(2):343–352
- Minsky ML (1986) The society of mind. Simon and Schuster Inc, NY
- Minsky ML (1992) Future of AI technology. *Toshiba Rev* 47(7):139
- Minsky M (2000) Commonsense based interfaces. *Commun ACM* 43(8):67–73
- Minsky ML (2006) The emotion machine: commonsense thinking, artificial intelligence, and the future of the human mind. Simon and Schuster Inc, NY
- Minsky M (1975) A framework for representing knowledge. In: The psychology of computer vision, McGraw-Hill, NY, pp 211–277
- Modha DS, Ananthanarayanan R, Esser SK, Ndirango A, Sherbondy AJ, Singh R (2011) Cognitive computing. *Commun ACM* 54(8):62–71
- Morgan B (2013) A substrate for accountable layered systems. PhD thesis, Massachusetts Institute of Technology
- Morgan B (2010) Funk2: a distributed processing language for reflective tracing of a large critic-selector cognitive architecture. In: proc. fourth IEEE international conference on self-adaptive and self-organizing systems workshop (SASOW), IEEE Computer Society, CA, pp 269–274
- von Neumann J (2012) The computer and the brain, 3rd edn. Yale University Press, New Haven and London
- Pal SK, Banerjee R (2013) Context-granulation and subjective information quantification. *Theor Comput Sci* 448:2–14
- Pal SK, Banerjee R, Dutta S, Sen Sarma S (2013) An insight into the Z-number approach to CWW. *Fundam Inform* 124(1–2):197–229
- Payne SJ, Reader WR (2006) Constructing structure maps of multiple on-line texts. *Int J Hum Comput Stud* 64(5):461–474
- Picard R (1997) Affective computing. MIT Press, MA
- Pinker S (1997) How the mind works. W. W. Norton & Company, NY
- Pinker S (2007) The stuff of thought: language as a window into human nature. Penguin Books (Viking Press), NY, USA
- Price CJ (2000) The anatomy of language: contributions from functional neuroimaging. *J Anat* 197:335–359
- Ramachandran VS, Blakeslee S (1999) Phantoms in the brain: probing the mysteries of the human mind. William Morrow and Company (Harper Collins), New York
- Ramachandran VS, Hubbard EM (2001) Neural cross wiring and synesthesia. *J Vis* 1(3):67
- Ramachandran VS, Hubbard EM (2003) The phenomenology of synaesthesia. *J Conscious Stud* 10(8):49–57
- Robinson K, Aronica L (2013) Finding your element: how to discover your talents and passions and transform your life. Viking (Penguin Group), NY
- Roese NJ (1997) Counterfactual thinking. *Psychol Bull* 121(1):133–148
- Rothkopf EZ (1971) Incidental memory for location of information in text. *J Verbal Learn Verbal Behav* 10(6):608–613
- Rugg MD, Yonelinas AP (2003) Human recognition memory: a cognitive neuroscience perspective. *Trends Cogn Sci* 7(7):313–319
- Ryle G (1949) The concept of mind. University of Chicago Press, USA
- Seth AK (2010) The grand challenge of consciousness (opinion article). *Front Psychol* 1(5):1–2
- Seth AK, Izhikevich E, Reeke GN, Edelman GM (2006) Theories and measures of consciousness: an extended framework. *Proc Natl Acad Sci (PNAS)* 103(28):10799–10804
- Singh P (2003b) Examining the society of mind. *Comput Inform* 22(6):521–543
- Singh P, Barry B, Liu H (2004a) Teaching machines about everyday life. *BT Technol J* 22(4):227–240
- Singh P, Minsky ML (2004) An architecture for cognitive diversity. In: Davis D (ed) Visions of mind. Idea Group Inc., London
- Singh P, Minsky M, Eslick I (2004b) Computing commonsense. *BT Technol J* 22(4):201–210
- Singh P (2003) A preliminary collection of reflective critics for layered agent architectures. In: Proceedings of the safe agents workshop (AAMAS), Melbourne, Australia
- Singh P (2005) EM-ONE: an architecture for reflective commonsense thinking. PhD thesis, Massachusetts Institute of Technology
- Singh P, Minsky ML (2003) An architecture for combining ways to think. In: Proceedings of the international conference of the integration of knowledge intensive multi-agent systems, pp 669–674
- Sloman A (1978) The computer revolution in philosophy: philosophy, science and models of mind. The Harvester Press Ltd., Sussex
- Sloman A (1984) Towards a computational theory of mind. In: Artificial intelligence—human effects, Ellis Horwood, UK, pp 173–182
- Sloman A (2001) Varieties of affect and the CogAff architecture schema. In: Proceedings symposium on emotion, cognition, and affective computing AISB’01 convention, pp 39–48
- Snaider J, McCall R, Franklin S (2011) The LIDA framework as a general tool for AGI. In: Schmidhuber J, Thórisson KR, Looks M (eds) Proceedings of 4th international conference in artificial intelligence, vol 6830 of lecture notes in computer science, Springer, pp 133–142
- Stallman RM, Sussman GJ (1977) Forward reasoning and dependency-directed backtracking in a system for computer-aided circuit analysis. *Artif Intell* 9:135–196

- Stocco A, Lebiere C, Anderson JR (2010) Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol Rev* 117(2):541–574
- Sussman GJ (1973) A computational model of skill acquisition. PhD thesis, Massachusetts Institute of Technology
- SyNAPSE. <https://www.research.ibm.com/cognitive-computing/neurosynaptic-chips.shtml>. Accessed 8 April 2014
- Todorovic D (2008) Gestalt principles. *Scholarpedia* 3(12):5345
- Turing AM (1950) Computing machinery and intelligence. *Mind* 49:433–460
- Turing A (1949) Intelligent machinery. [http://www.alanturing.net/intelligent\\_machinery/](http://www.alanturing.net/intelligent_machinery/)
- Wertheimer M (1923) Laws of organization in perceptual forms. *Psychologische Forsch* 4:301–350
- Winograd E (1988) Some observations on prospective remembering. In: *Practical aspects of memory: current research and issues*, vol 1. John Wiley, NJ, pp 348–353
- Winograd T (1971) Procedures as a representation of data in a computer program for understanding natural language. PhD thesis, Massachusetts Institute of Technology
- Winston PH (1970) Learning structural descriptions from examples. PhD thesis, MIT
- Wolf M (2007) Proust and the squid: the story and science of the reading brain. Harper Collins, NY
- Zadeh LA (1994) Fuzzy logic, neural networks and soft computing. *Commun ACM* 37(3):77–84
- Zadeh LA (1996) Fuzzy logic = computing with words. *IEEE Trans Fuzzy Syst* 4(2):103–111
- Zadeh LA (1998) Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. *Soft Comput* 2:23–25
- Zadeh LA (2011) A note on Z-numbers. *Inf Sci* 181(14):2923–2932
- Zhang Z, Franklin S, Dasgupta D (1998) Metacognition in software agents using classifier systems. In: Mostow J, Rich C (eds) *Proceedings of fifteenth national conference on artificial intelligence and tenth innovative applications of artificial intelligence conference*, AAAI Press, CA, pp 83–88