Z* Numbers Revisited: Extraction and Consolidation for Natural Language

Kumar Kshitij Patel[1], Romi Banerjee[2], Sankar K. Pal[3], Life Fellow, IEEE

***Abstract*— The human ability to use language and communicate with fellow beings is arguably the cornerstone of modern civilization. This research is an endeavor to understand some elements of this process and simultaneously use this understanding to enable a machine-mind framework to understand natural language. Z* numbers serve as tokens of mentalese, forming a medium where thoughts can nurture. It is crucial for them to extract information from natural language, while enduring minimum loss. Also, various operations need to be defined over the Z* numbers so that they can be used in diverse computational and cognitive tasks. In this article, we propose an agency based framework inspired by Marvin Minsky's Society of Mind, describing multiple building blocks for the tasks of extraction and consolidation. The act of consolidation not only reduces memory required to store the information, but also forms the basis of all understanding tasks, by filtering out the irrelevant linguistic matter. These agencies would use basic cognitive resources and knowledge structures. While only basic ideas are provided for the latter, a robust mapping is done between agencies and cognitive resources, characterizing how they work. The elements of a Z* number are elaborated further while describing a functional ontology, that aids their computational usage. Some new ideas about them are also presented while preserving the basic intuition behind them. Owing to this bottom up approach, the complex task in hand is simplified and assigned simpler subtasks. We also offer a sample run-through of the architecture to better explain its working. This follows some real world experimentation where we try to observe analogous processes in humans and draw insights into the working of our society of agents. This research not just extends our own machine mind framework, but also aids general cognitive architecture, natural language understanding and hybrid knowledge systems research.**

***Index Terms:* embodied cognition, generally intelligent man-machine systems, natural language understanding, self-conscious systems, thinking machines, cognitive architectures, society of mind, hybrid knowledge systems.**

I. INTRODUCTION

WITH his seminal paper on machine consciousness[1], one of the greatest visionaries of the Human Colossus, Alan Turing gave humanity a new dream. A dream it would pursue for decades to come. In the process, loosing faith and even dreading its possible consequences at times. The dream and the following pursuit of artificial intelligence has effected research and industry for more than 60 years now. After the "second AI winter" and the "information revolution", there have been multiple path breaking events in AI research. With the rise of machine learning and other statistical tools we have made tremendous progress in specific intelligence tasks. Many of these abilities have not only equaled but surpassed the human counterparts. However, there hasn't been comparable investment and development in general intelligence. Though, the cognitive architectures have developed immensely since Kasparov lost the game; merging specific intelligence algorithms, optimization techniques, knowledge representation and powerful hardware, we haven't been able to do tasks which humans do without thinking in their daily lives. [2] Presents a brief survey of cognitive architectures while mentioning possible research goals to drive the industry. From the survey and otherwise it is very clear that architectures in hope of capturing one cognitive element have consistently failed on others. Marvin Minsky, one of the fathers of artificial intelligence in his two books[3][4], maintains the idea that a complex treatment of brain and hence machine-mind is inevitable owing to its inherent and inseparable complexity.

Our research which started with proposing a machine-mind framework[5], followed by the description of Z*-numbers[6] which are tokens of mentalese to data structures for endogenous thought[7], heavily borrows from the ideas developed in psychology, sociology, philosophy, linguistics, computer science and mathematics. While the long-term goal of this research is multi-modal perception and intelligence, we focus herein on only natural language. The association of linguistic abilities and the overall human intelligence has been studied for a long time now. Many describe the problem of language understanding and general intelligence to be one and the same for machines[8][9]. In fact researchers have gone ahead an described natural understanding as an AI-complete problem[8][9]. On the other hand, many philosophies in linguistics and pragmatics, like the one below by Mahavira (even in theology) acknowledge the inherent flaws in a language rendering it incapable to express complete reality:

> *Reality can be experienced, but it is not possible to totally express it with language. Human attempts to communicate is Naya, or "partial expression of the truth". Language is not Truth, but a means and attempt to express Truth. From Truth, language returns and not the other way around.*
> -Mahavira[10][11]

This research was conducted while Kumar Kshitij Patel was interning at the Center for Soft Computing Research, Indian Statistical Institute Kolkata, along with Romi Banerjee under the guidance of Dr. Sankar K. Pal.

[1]Kumar Kshitij Patel is a Junior Undergraduate at the Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur, kishinmh@cse.iitk.ac.in

[2]Romi Banerjee is a Senior Research Fellow at CSCR, ISI Kolkata, rm.banerjee@gmail.com

[3]Dr. Sankar K. Pal is a Life Fellow, IEEE and a distinguished scientist with CSCR, ISI Kolkata, sankar@isical.ac.in

While, we don't take sides we acknowledge that the mentalese translation to language can be very lossy. Z*-numbers which aim to represent this language, however can overcome this limitation by the virtue of the fact that they are designed for mentalese. They stand closer to the nature of thought than language or any individual form of sensory perception.

However, we don't live in a world where such communication even seems like a viable possibility. Thus it makes sense in the least to talk about a robust process of extracting Z* numbers from natural language. This isn't a task reducible to NLP because it is unnatural for a hybrid architecture like ours to use indirect statistical measures to understand language. The framework suggested within tries to use the complex cognitive tools introduced in [5], to the fullest capacities while also interacting with the knowledge representations; very similar to human minds. It is also clear to see that the task is not trivial owing to the indirectness and complexity of human language, even far fetched is the dream of making trivial mapping between language and Z*-elements. We try to proceed towards the far fetched dream here though.

Another inevitable task is operating on Z* numbers once they are extracted. One of the important operation is consolidation. It aims at combining the knowledge from multiple Z* numbers into single Z* numbers which can work as information granules. This is how we learn as humans too and form in order experience, memory, intuition, knowledge and ultimately wisdom. Each of these are basically such consolidations collected over lifetime. These consolidated granules form concepts and the context itself. This task is again not trivial, because knowledge combines in very non-intuitive and complex ways, owing to what we already know and what separates us as an individual.

These tasks of extraction and consolidation have been dealt with in a society of mind like framework[3], with multiple agents and agencies involved in every task or sub-task. Knowledge structures have been used and hinted upon at multiple places without a full characterization. They form a part of our future research. After giving a description of the agents involved, a mapping has been provided between the agencies and the higher cognitive resources while fully characterizing the control flow in architecture. This process not only links our work to the previous work but is another evidence of how the architecture is capable of supporting such difficult processes. A mathematical treatment is provided wherever ideas have gained sufficient ground for description, while at other places only partial operationalization has been done. These parts would also be a part of our future research.

We begin by describing a theory of agents, some aspects of language, context, culture and some of our previous work in the form of Z* numbers and the machine-mind framework. This is the lesser contribution of the paper but can be used by anyone curious to understand the underlying concepts in fields explained. In that sense it is a survey of some existing technologies too. We then proceed towards the original contribution where we first talk about the components of Z* numbers and give some new insights about them, it follows a description of general and specific agencies for the above two tasks, the mapping with the higher resources is laid down and supported with some clues from other sciences. Then, we present a run-through of the entire framework using some examples. A human experiment has also been done to elucidate the formal procedure and find useful insights.

Our Contribution: Novelty of this research lies in identifying and realizing micro and macro level agents involved in primary language understanding, linking them to higher cognitive resources including memory and formalizing the entire process while describing the ontologies and mathematical formulations. It compliments multiple research domains:

- *Natural Language Understanding:* Insights from the agency framework can guide more complex algorithmic styles. It not only captures the essence of hybrid architectures but merges them with centuries of psychological and neuroscience research.
- *Knowledge Based Systems:* Often criticized for their rigidity and lack of automation, this hybrid scheme gives insight on how they can be used along with more intelligent cognition inspired tools to enhance the process of reasoning and add the touch of human subjectivity to it.
- *Cognitive Architectures:* This research is a living example of how cognitive architectures grow overtime. It adds the modular-design factor to the architectures while not compromising with its own complexity. We aim to reduce the limitations of general intelligence systems and introduce the trend of combining multiple research results from all domains into architectures. While not a completely functional architecture yet, even in its infancy it promises the ability to deal with complex research problems like self reflection.

II. THEORY

A. *FROM SOCIETY OF MIND TO THE EMOTION MACHINE*

In his seminal book, Society of Mind[3] originally published in 1986 and followed by an equally enticing read The Emotion Machine[4] which came in 2006, Marvin Minsky describes a lifetime of understanding of human brain while laying down the foundation of one of the most protracted cognitive architectures. It is one of the few architectural endeavors that aims to capture the human brain in its full complexity without simplifying concepts and processes. It acknowledges its complexity and lays down a bottom up approach at building 'agents' for specific tasks. Some of his findings and suggestions that have been used in this research are[5]:

- *Agents* are the building blocks of any mind architecture. They represent simple tasks which usually are simple

mappings to realizable processes or actions. They are easy to 'understand' as compared to the complex tasks they try to achieve in coalition. [12] describes an agent as a generalized complex granule with inbuilt control mechanisms.
- *Agencies* are societies of agents that in totality perform tasks more complex than a single agent. An agent doesn't know his job while an agency does as he describes. For e.g. a circuit is an agent while a machine is an agency.
- *K-line* is a special type of agent which turns on other agents. These are of two types *Nemes* and *Nomes*. Nemes are responsible for representation of an idea or context in mind. Nomes on the other hand control the very representation and affect agencies in a predetermined way. Some of them are explained below:
  - *Polyneme:* They 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.
  - *Microneme:* They bestow 'global' contextual signals to agencies all across the brain and handle subtle elements which cannot be crisply defined or lack specific terminology of situations. *"...those inner mental context clues that shade our minds' activities in ways we can rarely express..."*
  - *Isonome:* They 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. *"..The power of polynemes stems from how they learn to arouse many different processes at once, while isonomes draw their power from exploiting abilities that are already common to many agencies..."*
  - *Pronome:* Control the attachment of terminals to frames and are typically associated with the short-term memory representation of a particular role of an element.
  - *Paranome:* Operate on agencies across multiple mental realms simultaneously with identical effects across all of them. *"...certain pronomes can operate in several different realms at once. Let's call them paranomes to emphasize their parallel activities...by using these cross-connecting polynemes and paranomes, the activity in each realm can proceed sometimes independently, yet at other times influence and be affected by what happens in the other realms..."*
- *Frames* are a form of knowledge representation which formally describe a structure for a concept. Hierarchies of frames help in representing complex concepts and sub-concepts. Frames contain attributes which are assigned values linking them to other frames and agencies. Along with K-lines they are used in a majority of cognitive tasks. The idea of a frame is of historical significance too because of inspiring later object oriented knowledge representation.
- A total state of mind is a list that specifies which agents are active and which are quiet at a certain moment. A partial state of mind merely specifies that certain agents are active but does not say which other agents are quiet. This applies to the actions we are considering here too. We completely characterize them by specifying the active agencies.
- We learn by attaching agents to K-lines, but we don't attach them all with equal firmness. Instead, we make strong connections at a certain level of detail, but we make weaker connections at higher and lower levels. This would explain why its possible to choose from a set of actions when we have the appropriate register or memory.
- *Level-bands:* The connections described above occur in level-bands. The central level-band helps us in finding general resemblances between remembered events and present circumstances. The lower fringe supplies additional details. We use them only by default when actual details are not supplied. Similarly, the upper fringe recalls to mind some memories of previous goals, but again, we're not forced to use them except by default, when present circumstances do not impose more compelling goals.
- The lower levels represent objective details of reality; the upper levels represent our subjective concerns with goals and intentions.
- *Difference-Engines* are 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* are highly evolved agencies that yield initial behavioral solutions to basic problems like locomotion, defense mechanisms etc.
- *Types of Learning*

  I suspect that genius needs one thing more: in order to accumulate outstanding qualities, one needs unusually effective ways to learn.
  -Minsky

  - *Accumulating:* Remember every experience as a separate case.
  - *Uniframing:* Find a general description for multiple examples. *"...Our different worlds of ends and means don't usually match up very well. So when we find a useful, compact uniframe in one such world, it often corresponds to an accumulation in our other worlds..."*
  - *Transframing:* Form an analogy or mapping between two representations or *"...bridging between structures and functions or actions..."*
  - *Reformulation:* Find new schemes of representing

existing knowledge or *"..modifying a description's character.."*

  - *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.

- *The Investment Principle:* Our oldest ideas have unfair advantages over those that come later. The earlier we learn a skill, the more methods we can acquire for using it. Each new idea must then compete against the larger mass of skills the old ideas have accumulated.
- Memories are processes that make some of our agents act in much the same ways they did at various times in the past.
- *The Recursion Principle:* When a problem splits into smaller parts, then unless one can apply the mind's full power to each subjob, one's intellect will get dispersed and leave less cleverness for each new task.
- *Uniframing* means combining several descriptions into one while observing some similar characteristics in them. A simple theory of when we start new uniframes would be that in the brain, there is an architectural constraint on how many K-lines are directly accessible to various types of agents. Thus, beyond the limit we start making a common frame. Consolidation operation would give us insights on how uniframing can occur.
- *The Exception Principle:* It rarely pays to tamper with a rule that nearly always works. It's better just to complement it with an accumulation of specific exceptions.
- *Thinking:* [13][14] Propose a critic-selector model of thinking for the human mind. In this model the machine identifies an optimal thinking strategy by "thinking of thinking" and then proceeds to solve the problem in hand using other agencies and its own knowledge representation and reasoning.
- *Six-layered model of mind:* Fig. 1. shows the six layered structure of mind proposed by Minsky. Its various components are explained as[5]:

  - *Instinctive Reactions:* These are evolutionarily coded instincts that aid in survival. They constitute a procedural and somewhat declarative knowledge base, to solve daily life problems and help in predestined learning too.
  - *Learned Reactions:* These are again procedural databases which have been reinforced positively or negatively over experiences, specifying the best solution/strategy to any problem in hand.
  - *Deliberative Thinking:* This involves consideration of several alternative solution approaches, and choosing the best; using logic and commonsense reasoning to select solution paths.

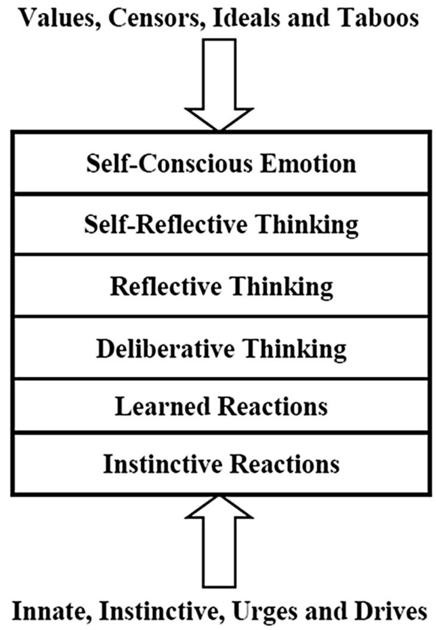


Fig. 1. Six-Layered Structure of The Mind[?]

  - *Reflective Thinking:* This involves self-introspection over solution strategies that came handy/failed pertaining to a problem solution. They are ranked and reinforced after every incident/thinking episode updating the knowledge base.
  - *Self-Reflective Thinking:* This involves reflection on oneself as a thinker. While the reflective layer focuses on the recent thoughts, it focuses on the entire process of thinking itself.
  - *Self-Conscious Emotion:* Verification of accordance of decisions with ideals, include self-appraisal by comparing ones abilities with others.

We'd be using the agent-agency framework in our design. The ideas and principles above form design directives which simplify a lot of our thinking.

B. Z-NUMBERS*

Following the introduction of Z-numbers by Zadeh in 2011[15], Banerjee et al. introduced Z*-numbers in 2015[6][16] to capture the human subjectivity. Given a natural language sentence Y on a subject X, its Z* number is a 5-tuple given by $\langle$T,C,A,B,AG$\rangle$ where:

- T is the time- a combination of past, present and future moments,
- C is the context associated with Y,
- A is the value of X i.e. an instantiation (or, restriction on the values) of X given C and T usually derived from the predicate of Y,
- B is a measure of reliability or certainty for A given X, C and T and
- AG stands for affect groups-a set of ordered triples of (affect,affect-strength,affect-valence)-arising on account

of mind transactions triggered by X given A,B,C,T. Valence belongs to {+, -}, while strength uses different intensity keywords which can be described as fuzzy classes. Affect types arise from the basic 6 emotions described in [17]. The triplet is capable of expressing a variety of mixed emotions.

A and B are perception-based fuzzy numbers while, AG is similar to spectral fuzzy numbers[18]. Z*-valuation is a 6-tuple given by $\langle$X,T,C,A,B,AG$\rangle$. Some preliminary properties of Z* numbers as proposed in the original paper are:

- The value of A, B and AG need to be precisiated using membership functions $\mu_A, \mu_B and \mu_{AG}$ respectively.
- μ_{AG} uses as operands the polarity and weights of the affects constituting AG. The sub affects of AG are also fuzzy numbers and need to be precisiated using membership functions.
- X and A together define a random event given (T,C) in R, and the probability p of this event is given as

$$p = \int_R \mu_A(u) p_X(u) du$$

where p_X is the underlying probability density of x and given (T,C), u is a generic value of X.

- Z*-valuation is considered a generalized constraint[?] on X and is defined as-

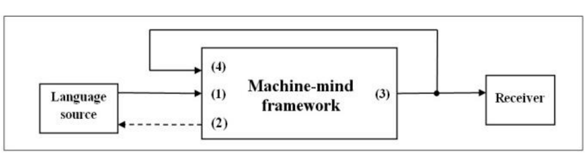
$$Probability(X\ is\ A)\ is\ (B *^C AG)\ given\ (T, C)$$

$$or, p\ =\ \int_R \mu_A(u) p_X(u) du\ is\ (B *^C AG)$$

where $(*^C)$ denotes a complex operation involving B and the constituent affects in AG to agglomerate the mutual roles of belief and emotions indecision-making. The operator wasn't characterized in the original paper and we propose an analysis here.

Z* numbers once computed, along with the machine mind framework[5] can be used to model real life conversations[6] where they not only aid in a better subjective representation but also make reasoning simple. For a running example see[6].

C. *THE MACHINE MIND FRAMEWORK*



Index –

(1) – System input => multi-modal language-elements.

(2) – Re-processing requests ('re-read text section', 're-listen sample portion', etc) due to incorrect or insufficient comprehension.

(3) – System output = 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 system 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 language-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 to the current processing activities.

Fig. 2. Incremental-developmental strategy of comprehension

Proposed in [5][16], the machine mind architecture has been shown in Fig. 3. As recognized in [5], the relevant tasks in designing a machine mind for text comprehension are:

1) Identification of the basic operations of the mind during text understanding.
2) Segregation of the operations into broad categories (or 'agencies').
3) Enumeration of the fine-grained 'agents' that underlie the agency-operations.
4) Construction of the elements of intra-agency and inter-agency communication and agent-activation.
5) Designing a model architecture that supports all of the above.

(1), (2) and (5) were dealt with in [5] (Summarized below), (4) was dealt with in [6][16] (Z*-Numbers) and this paper is an attempt to describe some details surrounding (3).

Before we study the agencies involved in the machine-mind, we briefly enumerate the processes involved in natural language understanding for the same. Fig. 2. illustrates the incremental-developmental strategy for forming comprehension while it follows the following steps using multi-sensory input and premeditated knowledge about the world:

- *Prediction:* Causally relate the present to past experiences and visualize future actions on the basis of intuition, common-sense, reinforced learning and reflection.
- *Visualization:* Conjure mind-images (real or intentional[19]) depicting people, places, events, etc., from natural language components.
- *Connection:* Build inter-domain associations, and existing and new knowledge.
- *Question and Classification:* Reflect upon and test strength, completeness, correctness and relevance of knowledge associations; re-organization and rectification of associations.
- *Evaluation:* Test coherence between perception granules formed for the current processing event, measure relevance and prune insignificant granules; attach notions of subjectivity or 'self-consciousness' (emotions, degrees of interest, summarize, biases, etc.).

The processes that underlie these complex functions are[5]:

- *Symbol extraction and symbol generation:* 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' in-

terpretations of unfamiliar words and phrases from surrounding text and the genre. Some of its subtasks are:

  - *Anaphora/Cataphora resolution:* Antecedent or subsequent context reliance among sentences/phrases is extracted and utilized in other tasks.
  - *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 homonyms 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 evolution:* 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 operations:* Creation, recall and operate upon frames and frame-systems to form concept granules across different level of granularity. *"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 frames terminals, consistent with the context sensitive assignment-conditions, system expectations or surprises and the envisioned system goal"*[5]. The different kinds of frames for linguistic entities are:
  - *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.
- *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 affect affect 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[13][14].

As shown in Fig. 3., following the society of mind kind of architecture we have divided our framework into super-agencies which have clusters of sub-agencies. Each of the super-agencies handle complex tasks like sense evaluation, reasoning, control etc. The super-agencies and their components have been explained below:

- *Sensory-Gateway:* It acts as the receiver for all sensory signals. It has modalities of Vision(**V**), Audition(**A**), Olfaction(**O**), Tactile(**Tc**), Taste(**Ta**), Balance(**B**), Temperature(**Te**), Pain(**P**) and Kinesthetic(**K**). For the purpose of text comprehension only vision is relevant. It extracts words, does primary morphemic analysis and references **L** and **ComN** to extract meanings of morphemes through **Log** and **De**. Error handling occurs here for stray marks etc. System results can be conveyed to the external world as shown in the Fig. 3. Other features associated with reading, like saccade length, speed, time and location of retrievals etc are also controlled by this super-agency. The local b of the sub-agencies under **SG** is analogous to the sensory memory concept in the human brain.
- *Deducer:* 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 under De use their local **FA** workspace to reason through the applicability of multiple solution-perspectives before globally advocating (a $\langle problem, solution, reason \rangle$ tuple) frame manipulation processes through Log. The sub-agencies of interest are:
  - *Syntax:* It is responsible for syntax-resolution and consequent generation, manipulation, pruning of surface syntactic frames. Activates relevant **ComN**

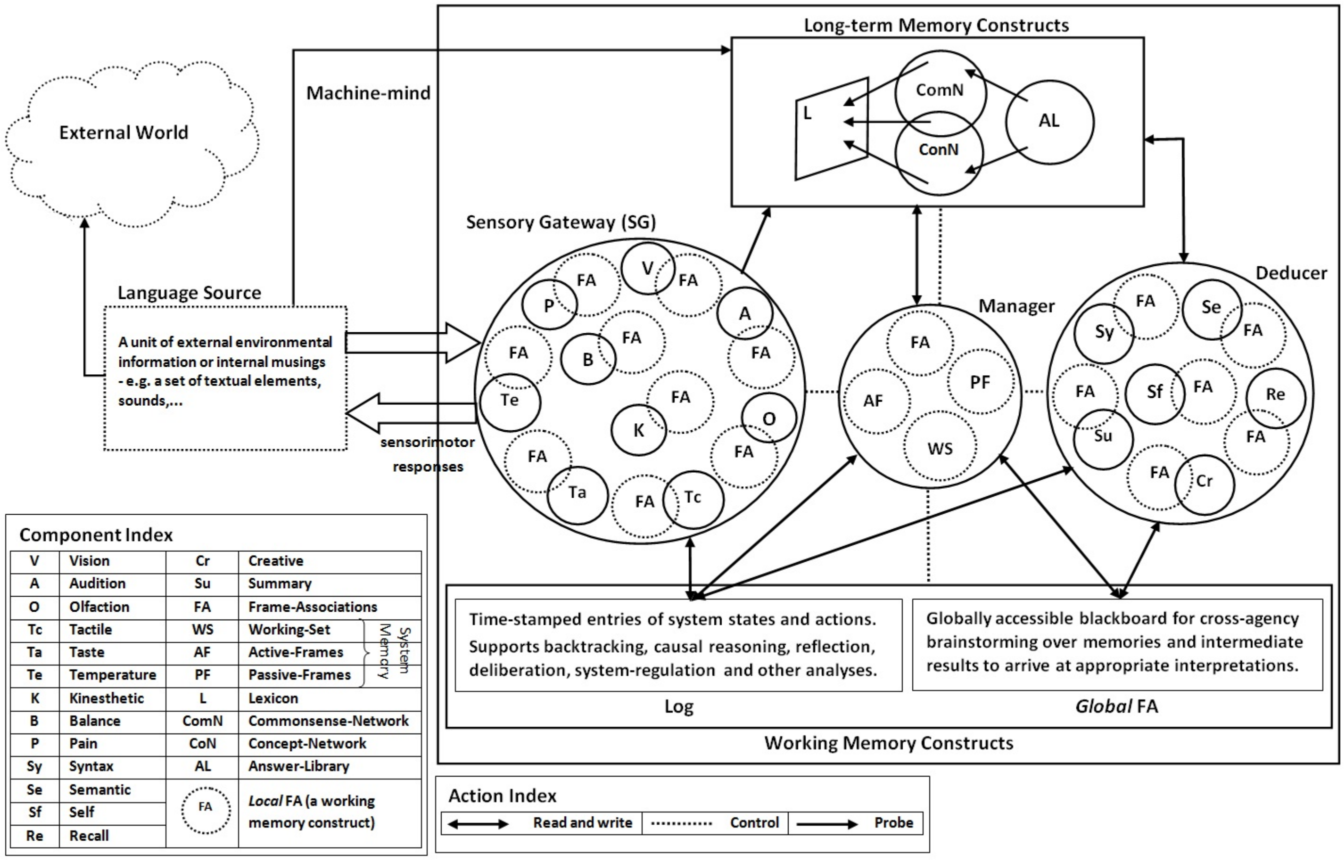


Fig. 3. Machine Mind Framework[5]

and **CoN** sections.

- *Semantic:* It is responsible for semantic-resolution of the text unit being processed after realizing its genre and generation, updating of surface semantic, narrative and thematic frames. Activates relevant **ComN** and **CoN** sections.
- *Self:* It flavors comprehension with components (affects, biases, ideals etc.) of the system self and activates multi-realm thinking as follows:
  - *Monitors* Affect-progression, belief and confidence of knowledge received, attention progression, reinforcement of knowledge (over **CoN**, **ComN** and **AL**) by identification of and interaction with attachment figures or self assessment.
  - *Initiates* upgrading of heavily reinforced **L**, **AL** and **CoN** elements to **ComN** triggering predestined learning.
  - *Effects* recollection of memories; belief and confidence are associated measures.
  - *Manipulates* semantic, narrative and thematic frames.
  - *Ensures* cognitive biases do not lead to incorrect processing.
  - *Spawns* multi-mental realm reformulations of a problem; each realm in turn activates relevant agencies.
  - *Achieves* self-reflection by judging the alignment of the text to ideals and preferences.
- *Recall:* It thin-slices a problem into sub-problems, mapping problems to memories and retrieving the same from long-term memory for processing in the current context. If all sub-problems have known solutions, it activates memories of solutions in **AL** and initiates involvement of the required agencies in the text interpretation processes. For sub-problems that have no solutions, it activates **Cr**. In parallel, it also activates **Su** to monitor and conquer partial solutions to an effective mechanism while updating **AL**, **ComN** and **ConM**.
- *Creative:* It projects and suggests solutions for 'new' problems and is the hub of reflection, imagination, creativity and system IQ[?]. Its tasks are elaborated as follows:
  - *Hypothesizes* interpretation strategies for a given 'new' problem.
  - *Evaluates* differences between a problem and the 'similar' experiences recalled by Re. Reformulates, accumulates and uniframes memories. Transframes across contexts and memories. Using commonsense and intuitive reasoning, it improvises upon known 'similar' solution strategies to counter differences; initiates solution trials by

other sub-agencies.

* *Builds* solutions from scratch by initiating trans-framing trials and subsequent solution trials.
* *Handles* exceptions; 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.
* *Activates* **Su** to monitor solution trials to an effective mechanism.
* *Initiates* updating of **L**, **CoN**, **ComN** and **AL**.
* *Emulates* 'imagination' i.e. the ability to visualize intentional objects[19].

– *Summary:* It measures distance between the current state of the system and the projected goal through relevance, affect and comprehension progression evaluation resulting in activation or inhibition of agencies, while consolidating memories. Some of these processes are elaborated below:

* *Predicts, visualizes, questions* and *clarifies* all computational mind activities during text processing.
* *Monitors* relevance and comprehension- progression through text processing.
* *Generates* curiosity [22], questions in the computational mind, when comprehension is incomplete or unsatisfactory by measuring information gaps [23], attention and interest, to regulate saccade length and consequent text-intake rate by **V**. Also, instructs **V** to re-read or search for textual cues that relieve curiosity.
* *Adjudges* non-convergence of syntactic or semantic analyses and inhibits erroneous operations, leading to the identification of semantic errors in text.
* *Consolidates* solution principles of sub-problems (to formulate effective text-interpretation strategies) and frames resulting out of sub-problem solutions into coherent granules of facts and events.
* *Deliberates* and *reflects* over successful and unsuccessful interpretations and strategies used thereof to reason or clarify success and failure. Also, reflects over inhibited processes to emulate 'counterfactual' thinking. These Reflections motivate 'new' thinking by activating **Cr** which in turn triggers other sub-agencies.
* *Applies* these new interpretation procedures, formed by **Cr**, to problems ranked 'similar' by **Re** while comparing these strategies against old strategies and updating **AL**. It also annotates solutions with $\langle\ problem, process,\ result,\ reason\ \rangle$ for storage in **AL**.
* *Annotates* memories with $\langle\ environmentdescriptors,\ problem,\ solution, result,\ reason,\ affects,\ beliefs,\ etc.\rangle$ for storage in CoN.
* *Segments* text into sections: introduction, rising action, climax, resolution, and denouement, based on information, affect and interest progression.
* *Updates* **L**, **CoN**, **ComN** and **AL**. Updating of **AL** triggers upgrading of the agents that symbolize algorithms under sub-agencies.

Based on the information granules they use the sub-agencies under **De** can be classified as follows:

– *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**).

• *Manager:* It is the global administrator or 'heart' of the system. It runs in the background and is responsible for the activation and execution of 'involuntary' functions that support the functioning of all the other agencies. It does continual self-evaluation of system processes and updates towards improved (cost effective and robust) system performance. 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. Some of its tasks are:

– *System time management:* Maintains system clock for Log entry timestamps and ensures real-time time constraints over operations such that system cognition is at most of the order of average human cognition rates.
– *Attaches* unique identifiers to extracted saccadic information for use in **Su**.
– *Memory handling:* It has to monitor and direct inter and intra memory transfers between working and long-term memories while maintaining uniform encoding and avoiding thrashing in either of these. The system memory-management constructs used by **M** are:

* *Working-Set:* Set of pointers to frame-networks in **FA** being referenced within a narrow time window (intuitively, of the order of seconds).
* *Active-Frames:* 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**.
* *Passive-Frames:* 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.

- *FA management:* It maintains coherence across local and global **FA** while selectively pruning local **FA**, annotating 'trial' and 'applied' results and allotting fixed-size or adaptive physical memory space for local **FA**.
- *Log-management:* Read/write synchronization across multiple agencies, commit point handling (write-back all 'correct' short-term memory modifications to long-term memory constructs) and heuristic scheduling[25] to arbitrate multiple agency attention seeking requests.
- *Context-switching:* It involves storing the status of the current context and transferring control to a new context.
- *System optimization:* It utilizes idle processor cycles to perform on-line housekeeping tasks, reflects over system management mechanisms to reason and self-modify towards enhancement and executes **Su**s efforts to arrive at 'new revelations'.
- *K-Line Management:* It is responsible to create or kill a K-line component (like identifier-assignment, memory management, Log entries). Tasks undertaken by different K-line components in accordance with their description in "Society of Mind" or otherwise are:
  * *Polyneme:* Tracks FA components denoting different ideas about a singular parent-frame; every different sense of a homonym has a unique polyneme tracking its corresponding FA elements.
  * *Microneme:* Encodes global context parameters, as evaluated by **Se** which are 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.
  * *Paranome:* Tracks FA components pertaining to an active mental realm of thinking for the given text.

- *Long-term Memory Constructs:* This is the knowledge repository of the architecture. Its multiple components are:
  - *Lexicon:* System vocabulary (words, phrases, idioms) and their meanings encoded in machine 'understandable' form, either precise machine-language statements or multi-modal implications (sounds,images and metaphors).
  - *Answer-Library:* Resource of $\langle solution\_strategy,\ result,\ reasons\rangle$ for $\langle context\_parameters,\ problem\rangle$ tuples.
  - *Concept-Network:* Network of networks of inter-contextual concept granules, a hyper-graph of associations across frame-systems. Formed consciously or unconsciously but the elements are retrieved 'consciously'.
  - *Commonsense-Network:* 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.
- *Working Memory Constructs:* These are referenced by all the agencies and form the basis of deliberative and reflective actions of the system. Its parts are as follows:
  - *Log:* It is an on-line global record of time-stamped agency-activity entries. It indicates instantaneous state of the system, analyzing which agencies and agency functions may be autonomously or exogenously(by **De**) activated. It also initiates mechanisms like intelligent backtracking[24], generates error signals and serves as an indicator of solution strategy results and reasons thereof forming the basis of self-reflection.
  - *Global Frame-Association:* It is a blackboard or scratch-pad for frame-system manipulations during the process of text 'understanding'. 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. Also, all 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. Each sub-agency can share sections or all of its local **FA** with the other agencies, through Global **FA** or **Log**. 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.

Fig. 4. shows a functional summary of the architecture. For working, more insights on implementation and processing refer[5].

D. *THEORIES OF EMOTION*

Emotion has been extensively studied in affective computing, sentiment analysis and cognitive sciences. For the purpose of cognitive architecture the utility of a theory of emotion is twofold, it helps in understanding human behavior as well as in artificial emotion generation for a social robot. We survey[1]some prominent theories of emotion and identify the common characteristics they share. For the sake of brevity, we describe two frameworks of analysis here.

[1]Refer[26][27] for a fantastic survey on theories of emotion. This section borrows from these sources.

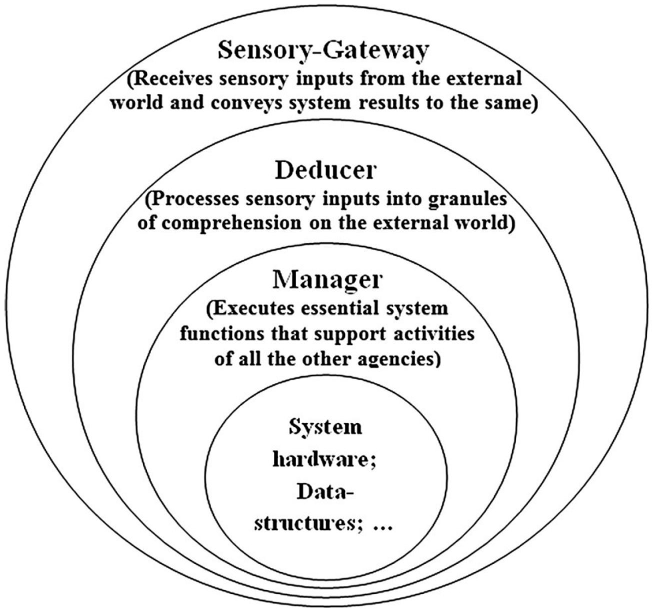


Fig. 4. A Diagrammatic Summary of Machine Mind[5]

TABLE I
COMPONENTS OF EMOTIONS AND EMOTIONAL ARCHITECTURES[26]

Component	Functions
Cognitive	Stimulus evaluation/Appraisal
Feeling	Monitoring → Regulation
Motivational	
Somatic	Preparation and support of action
Motor	Action

[26] Defines an emotional episode as *"...to indicate anything starting from the stimulus to the later components or the immediate consequences of the emotion. The notion of emotional episode is thus potentially broader than the notion of emotion..."*. These later components are as shown in table I[26]. The theories of emotion talk about some or all of these components while introducing causality and sequentiality. [26] further elaborates on the cognitive component:
"The meaning of the term cognition seems to shift depending on the category with which it is contrasted.....the cognitive component can be understood in the broad sense of mental or in the more narrow sense of non-dynamic, Intentional, propositional, non-automatic, or rule based"
Marr in [28] suggested that emotion theories can be talked about on three levels, functional, algorithmic and implementational. Besides causation, most theories converge on the fact that there is some quantity and quality to emotions. Thus, from a causation perspective we can talk about three things:

- Elicitation of emotions
- Their intensity (quantity)
- The associated differentiation (quality)

Further, each of these three topics can be studied in terms of internal and external factors i.e the associated stimulus and realizing mechanisms in human body. This analysis has been summarized in table II[26]. Different theories work on different levels, include different components, refer to different sequentiality and causality among them and select different combinations for constituting the emotional experience. Theories also differ in dimensionality and in whether they consider basic emotions as the building blocks or some sub-emotion modules that build up emotions further.

We use this framework as we describe some prominent emotion theories. We have left out some historically relevant theories and philosophical theories. Critical comments have also been avoided, (for a detailed survey visit [26][27]). The theories are as follows:

- *James' Theory:* It is considered a feeling theory because it equates the emotional experience with emotion. It is functional and slightly implementational in nature. The stimulus is supposed to act on the sensory system which directly cause the somatic peripheral changes in the body. The conscious realization and interpretation of these changes forms the emotional experience. Thus, the intensity and quality are both controlled by subsequent parts in the bodily responses. Each emotion has its own body signature. Elicitation is not explained. Damasio's and Prinz's theories are considered neo-Jamesian.
- *Scachter's Theory:* Developing on James' theory to tackle the criticism regarding intentionality and specificity of emotion, Scachter proposed this feeling theory. The stimulus leads to a physiological arousal, which follows its attribution. This attribution leads to the the emotional experience. Degree of somatic arousal decides intensity while the attribution decides differentiation. It again fails to cite the elicitation process.
- *Appraisal Theories:* It is one of the most discussed theories of our times and past. Beyond the fact that an evaluation or appraisal precedes the emotion, there are large debates among the the community regarding this theory. However in the most cited idea, stimulus is followed by an appraisal which develops an action tendency leading to physiological changes. These changes affect behavior and finally there is an attribution or labeling of emotion. The process from appraisal to behavioral changes constitutes the emotional experience. The emotion-antecedent appraisal and emotion-consequent attribution are not much different in the cognitive sense but in the functional sense, as former inputs stimulus while latter inputs the emotion itself. These processes can be overlapping, recurrent and are usually necessary for the emotional episode. The appraisal theories usually focus on functional and algorithmic levels. One of the heavily discussed models is the OCC model [30][29], which has been both implemented extensively in artificial systems and been criticized for its one-to-one mapping, proposal of exhaustive stimuli and limited focus. In order to deal with these difficulties modern appraisal models have come up with variables like goal relevance, goal congruence, coping potential, certainty

TABLE II

QUESTIONS THAT NEED TO BE ASKED TO EMOTIONAL THEORIES BASED ON MARR'S LEVEL OF ANALYSIS[26]

	Problems Related to Emotion Causation		
Marrs levels of process description	*Question 1: Elicitation*	*Question 2: Intensity*	*Question 3: Differentiation*
A. Functional level: Relation between input and output	Question 1A: Which stimuli elicit emotions and which do not? What are the conditions under which emotions are elicited	Question 2A: Which stimuli elicit weak versus strong emotions?	Question 3A: Which stimuli elicit positive versus negative emotions? (anger, fear, sadness, joy, etc.)
B. Algorithmic level: Mechanisms and format of representations (codes)	Question 1B: What are the mechanisms and representations that determine emotion elicitation?	Question 2B: What are the mechanisms and representations that determine the intensity of emotions?	Question 3B: What are the mechanisms and representations that determine the quality of emotions?
C. Implementational level: Neurological structures or routes	Question 1C: What is the neurological basis of emotion elicitation?	Question 2C: What is the neurological basis of emotion intensity?	Question 3C: What is the neurological basis of emotion differentiation?

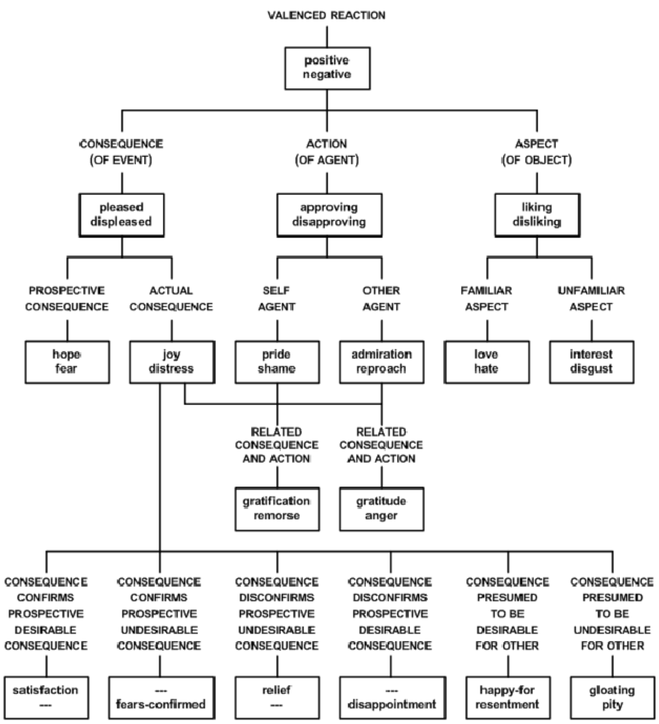


Fig. 5. OCC Model[29]

and agency/blame. Precise number of these appraisal variables is arguable. To develop more on the algorithmic side appraisal theories have come up with a multi-mode model. Rule-based, associative as well as sensory motor connections are proposed by these models for emotion elicitation. Since our model draws heavily from appraisal theories we present some practical motivations for the theory and some assumptions made to deal with them[29] in table III. OCC model has been shown in Fig. 5.[29].

- *Network Theories:* They draw inspiration from semantic network theories and are associative models i.e. they believe that emotions are recorded in memory networks. There can be two kinds of schemata, one evolutionarily hardwired and other acquired. The trigger for a schemata can be stimulus as well as the associated response or appraisal, that's why its difficult to draw sequential diagrams for these theories. While learning new networks there isn't any consensus whether space-time correlation is enough or not, but for retrieval theorists usually agree on an automatic, often unconscious mechanism which is in fact representation-mediated cognition. Clearly, they deal with algorithmic level and tend to ignore the functional level. They also tend to answer only the quantity question clearly. To meet with the limits of association, primarily the need for a robust stimulus-response mapping, they propose multi-mode models where rule based elicitation is allowed. Often the variables and components if used are interdependent and undergo recurrent updation so as to reach a stable state. The intensity is decided by the extent of activation of network.
- *Affect Program Theory:* This theory often without contradicting the previous theories talks about the par where evaluation of stimulus is converted to other components of emotion. The nature of this evaluation may be decided by any other theory. In essence, it talks at the implementation level once the evaluation is done. Each basic emotion is hypothesized to have a neural circuit (or signature) which has developed evolutionarily to serve adaptation. These circuits might be involved in any other bodily activity too. Once elicited beyond a threshold with specific input in absence of undesirable inputs the circuit runs so as to produce changes in other emotion components. This theory can easily align with any other theories above as it talks about the implementation level. The reason it is slightly different from the single-mode network theory is that the circuits

TABLE III

APPRAISAL THEORY MOTIVATIONS AND ASSUMPTIONS

Motivation	*Assumption Made for the Solution*
1. Differentiated nature of emotional response.	1. Emotions are differentiated by appraisals: each distinct emotion is elicited by a distinctive pattern of appraisal.
2. Individual and temporal differences in emotional response to the same event.	2. Differences in appraisal: different individuals who appraise the same situation in significantly different ways will feel different emotions; and a given individual who appraises the same situation in significantly different ways at different times will feel different emotions.
3. Wide range of situations that evoke the same emotion.	3. All situations to which the same appraisal pattern is assigned will evoke the same emotion.
4. Elicitation of the process of emotional response.	4. Appraisals start the emotion process, initiating the physiological, expressive, behavioral, and other changes that comprise the resultant emotional state.
5. The appropriateness of emotional responses to the situations in which they occur.	5. The appraisal process makes it likely that emotions will be appropriate responses to the situations in which they occur. Appraisal system has evolved to process information that predicts when particular emotional responses are likely to provide effective coping.
6. Irrational aspects of emotions.	6. Conflicting, involuntary, or inappropriate appraisal may account for irrational aspects of emotions; inability to control the motivational and perceptual bases of the appraisal process.
7. Developmentally and clinically induced changes in emotion.	7. Changes in appraisal: if theoretically specified appraisals of a situation change over the course of development or are changed by psychotherapeutic interventions, emotional responses to those situations should also change.

are hardwired and may not associate with the memories while elicitation occurs. Marvin Minsky in his book "The Emotion Machine" develops upon similar ideas. These theories are dedicated to the idea that basics emotions are the building blocks for emotional life.

- *Barrett's Conceptual Act Theory:* It is based on Russell's core affect theory[31]. He states that not emotions but sub-emotional variables of valence and arousal hold value. Their combination is ought to create "affective quality" for a stimulus which gives rise to "core affect" in a person. Core affect has a neuro-psychological and a mental side. Barrett however proposes that emotion still hold value and adds the feature of emotion categorization on these core affects via context and social conditioning. Thus emotions don't have any meaning with socio-cultural dynamics. Clearly, it is a two factor theory, stimuli elicit core affect and then the core affect is categorized using previous conceptual knowledge or "emotional scripts". Both parts can be multi-modal while constraining each other, but for latter emphasis is on association or rather constraint satisfaction. Every context is associated with an emotion script. "...Category representations are no propositional or static, but perceptual, embodied and situated..."[26]. Perceptual because of modal-specificity, embodied because of motor features and situated because of context dependence. Like network theories it talks mainly about quantity questions but categorization also answer the quality question to some extent. They are different than appraisal theory because emotions sit here on the boundary of society and individual and they are more important than appraisal itself. It is like a feeling theory in the sense that emotional experience has specificity.

E. ARTIFICIAL EMOTION SYSTEMS AND EMOTION RECOGNITION

Historically, in all the cognitive architectures there has been an attempt to realize and understand emotions. Power and Dalgleish suggested eight major questions about an emotion theory[27], as shown in table IV. Additionally, subsequent questions as proposed by [?] for an artificial emotion system have also been presented in the same table. These questions answer some surface queries about a theory and help us in understanding the emotion architecture of a machine too. We'd use this framework while suggesting an emotion theory for our cognitive architecture. [?] Also proposed a classification of all architectures according to these 10 questions.

Artificial emotions form a general intelligence problem. However, emotion recognition is a specific problem and it has been extensively studied in affective computing and machine learning. The models for emotion recognition can be rule-based, statistical or hybrid. In a general setting, rule based models should do syntax, phrase, word and sentence level analysis similar to the one proposed in [33]. However, rule based models tend to only look for keywords or emotionally functional words which include emotion keywords, modifiers, metaphors etc. There are some issues with this[34]:

- Ambiguity in association of keyword with the supposed meaning.
- Incapability to detect emotion expression at other levels like phrases.
- Loss of linguistic information in syntax and semantics. Often sarcasm and other complex expressions are neglected.

Statistical approaches usually use the bag of words model. This makes emotion keyword recognition difficult and often oversimplifies the emotion category[34] . Hybrid approaches

TABLE IV
QUESTIONS ASKED ABOUT AN EMOTION THEORY

Major Questions to Answer Regarding Emotion Theories	*Emotion Theory Classification Questions for Artificial Emotion Systems*
1. What distinguishes an emotion from a non-emotion?	1. What distinguishes an emotion from a non-emotion within the system?
2. What are the constituent parts of an emotion, or are emotions irreducible?	2. What are the constituent parts of an emotion, or are emotions irreducible within the system?
3. What distinguishes one emotion from another?	3. Is there more than one emotion being used? If so, what distinguishes one emotion from another within the system?
4. What is the process of having an emotional experience?	4. What is the process of having an emotional experience for the system? Is it constant across emotions, or does it change for each emotion?
5. Why do we have emotions?	5. Why does the system have emotions? Do all emotions serve the same purposes?
6. What is the relationship between emotional states, moods, and temperament?	6. How do the systems emotional states differ from how humans characterize moods and temperament? Are emotions transient? If so, how long do they persist and what causes them to change?
7. How many emotions are there and what is the nature of their relationship with each other?	7. How many emotions does the system have and what is the nature of their relationship with each other? Do some emotions cause changes in other emotional states?
8. What is the difference between, and the relationship of, the so-called normal emotions and the emotional disorders?	8. Does the system detect and correct for the difference between, and the relationship of, the so-called normal emotions and the emotional disorders? And, if so, how does it do this?
	9. Where do the systems emotions originate? Are they created explicitly, learned from the environment, learned from social interactions, or some combination of these?
	10. Does the emotion improve the systems performance, and if so, in what way?

mostly use Latent Semantic Analysis. On top of that, models like SVM are applied to learn emotional features. In recent years there have been some deep learning models too. [35] Proposed a hybrid knowledge based ANN, which has been used for in this task. Surprisingly, there hasn't been much usage of emotion theories and case-based reasoning in this task besides the emotion classification schemes like OCC which again tend to oversimplify the analysis. A systematic analysis is difficult to present owing to multiple strategies but the section on natural language processing would provide a better insight.

F. REVIEWING LANGUAGE

> *The limits of my language are the the limits of my world.*
> -Ludwig Wittgenstein

We consider language on the level of sentences[2].A sentence can be defined as "a grammatically self-contained speech unit consisting of a word, or a syntactically related group of words that expresses an assertion, a question, a command, a wish, or an exclamation, which in writing usually begins with a capital letter and ends with a period, question mark, or exclamation mark.[36]" Though the analysis of grammar might seem like an unnecessary task here, but it is necessary to formally define these concepts, for these have come out of hundreds of years of thinking about language. They not only create usable ontologies but make the treatment more intuitive and closer to reality. They must not be viewed as rigid rules but an attempt to understand the world around us. When people communicate four major intentions and hence functional sentences are usually observed:

- Declarative: To inform someone of something.
- Interrogative: To get information from someone.
- Imperative: To get someone to do something.
- Exclamatory: To express one's attitude about something.

Before attempting to parse them, we must reconcile the parts of speech in English grammar. In fact in any grammar, parts of speech are sets words which share similar grammatical properties:

- *Noun (names)* is a word or lexical item denoting any abstract or concrete entity; place, thing, idea, or quality.
- *Pronoun (replaces)* is a substitute for a noun or noun phrase. Pronouns make sentences shorter and clearer since they replace nouns.
- *Adjective (describes, limits)* is a modifier of a noun or pronoun. Adjectives make the meaning of another word more precise.
- *Verb (states action or being)* is a word denoting an action, occurrence, or state of being. Without a verb a group of words cannot be a clause or sentence. Form of verbs, called verbals, can serve as other parts of speech. Gerunds are verbals that act as nouns; they usually end in "-ing". Participles are verbals that act as adjectives; they usually end in "-ing" or "-ed".
- *Adverb (describes, limits)* is a modifier of an adjective, verb, or another adverb. Adverbs make writing more precise.

[2]This section borrows heavily from the book *English Sentence Analysis: An Introductory Course* by *Marjolijn Verspoor* and *Kim Sauter*, from *John Benjamins Publishing Company, Amsterdam*. For a detailed analysis refer the same.

- *Preposition (relates)* is a word that relates words to each other in a phrase or sentence and aids in syntactic context. Prepositions show the relationship between a noun or a pronoun with another word in the sentence.
- *Conjunction (connects)* is a syntactic connector; links words, phrases, or clauses. Conjunctions connect words or group of words
- *Interjection (expresses feelings and emotions)* is an emotional greeting or exclamation. Interjections express strong feelings and emotions.
- *Article or Determiner (describes, limits)* is a grammatical marker of definiteness (the) or indefiniteness (a, an). The article is not always listed among the parts of speech. It is considered by some grammarians to be a type of adjective or sometimes the term 'determiner' (a broader class) is used.

We'd describe the analysis of only declarative sentences, but other sentences with some modifications can also be analyzed in a similar fashion. A typical declarative sentence usually has the elements as described in Table V[36]. These sentence

TABLE V
ROLES AND FUNCTIONS OF SENTENCE CONSTITUENTS

Roles	Functions	Abbreviation
First participant	Subject	S
Process	Predicator	P
Something about the first participant	Subject attribute	SA
A second participant	Direct object	DO
Something about the second participant	Object attribute	OA
A third participant	Indirect object	IO
The setting	Adverbial	A

constituents often have a predictable order in the sentences. It turns out there are five such patters in English as usually observed:

- *The running pattern (Intransitive verbs):* S + P + (A). For this pattern you need a verb that expresses an action involving only one main participant. Some verbs that express a pure action are running, swimming, talking, cycling, listening, etc.
- *The being pattern (Copula verbs):* S + P + SA + (A). For the being pattern, you need a verb that does not have much meaning, but expresses the sense of the mathematical equal sign (=). The meaning of such a verb is merely to point out a link between the first participant and an attribute or a category. For instance, appear, grow, seem, look, be, make, smell, sound, become, prove, taste, feel, remain, turn, etc.
- The doing/seeing pattern (Mono-transitive verbs): S + P + DO + (A) . For this pattern, you need a verb that expresses an action or a (mental) experience such as perception involving two participants, one who does the acting or experiencing and one who is acted upon or perceived. There are many verbs like doing, for instance, holding, counting, building, kicking, and many verbs like seeing that express (mental) experience like feeling, hearing, believing, thinking, etc.
- The giving/buying pattern (Di-transitive verbs): S + P + DO + IO + (A). For this pattern, there must be an event involving at least three participants, a person who gives something to someone or does something for someone (the subject), then the thing that is given or done (the direct object), and the receiver (the indirect or benefactive object). Very few verbs can be used in such patterns. The most common ones are give, pass, send, tell, make, buy, and offer.
- The making/considering pattern (Complex-transitive verbs): S + P + DO + OA + (A). When used with this pattern, a verb like make has a sense of 'doing something' and thus causing the 'direct object' to belong to a new category. For e.g., wipe, drive, call, crown, name, or elect. A verb like consider, when used with this pattern, expresses that in the subjects mind the 'direct object' belongs to a certain category. For e.g., assume, prove, declare, certify, regard, or deem.

In grammar, a clause is the smallest grammatical unit that can express a complete proposition. There are two general types of clauses:

- *Independent* clauses form a meaningful unit by themselves.
- *Dependent* clauses cannot stand on their own because they function as a constituent of another clause. They are of following types:
  - *Adverbial* clauses act as verbs.
  - *Noun* clauses act as nouns in the sentence. They can function as S, O or SA.
  - *Adjective* clauses modify or describe a noun.

Based on the number of independent and dependent clauses in a sentence four structural types are possible:

- *Simple:* They consist of only one independent clause.
- *Compound:* They consist of two or more independent clauses. The are joined using conjunction and often have a fixed order.
- *Complex:* They contain at least one full dependent clause with its own subject and predicate.
- *Compound-complex:* They contain at least two independent clauses and at least one full dependent clause.

As evident a sentence can consist of many clauses. Each of these clauses in turn consists of phrases, which are either single words or grammatically ordered groups of related words that together function as a unit. Phrases form a separate subject of study and we'd not dwell into them here. A possible apprehension at this moment is that such analysis of grammar and further usage in a machine mind might render it incapable of understanding subjective forms of human expression that often ignore grammatical rules. This issue has been separately discussed in the end.

G. ADVANCEMENTS IN NATURAL LANGUAGE PROCESSING

Natural Language Processing as a multidisciplinary discipline has witnessed 60 years of development since late 1940s. Meanwhile, NLP research has become exceedingly

structured, simultaneously incorporating developments in linguistics and machine learning. After the statistical revolution in late 80s and 90s, followed by introduction of neural networks, most of the NLP tasks have moved from rule based to statistical solutions. However, in the past decade we have seen multiple developments which make use of knowledge base and/or grammatical rules along with statistical procedures to offer state of art solutions solving problems like "hapax legomenon" and nonsensical inferences. The general procedure of language processing hasn't changed much, primarily because of the inherent structure of languages. We give a brief introduction to the modern NLP procedure without giving much details about the associated algorithms and solutions. For a detailed survey refer[37].
The process begins by obtaining a set of sentences. For the case of English, unlike languages like Mandarin, sentence and word endings are clear. They are marked by whitespace and punctuation symbols. Once obtained they go through a series of transformations and analysis as described below:

- *Lexical Analysis:* This deals with either word level structures or smaller. Lexicon forms an important part of any knowledge base. Words can be used in multiple forms owing to inflection in languages. This poses many storage and computational issues which need to be sorted. Some basic processes involved here are:
  - *Word Recognition:* As mentioned above it is trivial for the case of English owing to white spaces. In more complicated languages this however poses a problem.
  - *Morphological Analysis:* It studies the pattern of word formation in language due to inflection, derivation and composition. Inflection talks about the affix addition to the root of the word. It is the only factor under consideration here. Some subtasks are:
    - *Lexicon Formation:* Owing to inflection there are basically two possible forms of storage. Either storing all forms separately (Item-Arrangement) such that every word has multiple morphemes or storing according to grammatical classes while specifying rules of derivation separately (Item-Process). Former requires impractical amount of memory usually while the latter is difficult to characterize fully along with exceptions.
    - *Combinatorics:* In case the latter storage strategy is used, it lays down the rules for morpheme-stem combination. Inflection (like single to plural) and derivative (like noun to verb) morphologies are the two rules considered here.
    - *Orthographic Analysis:* This lays rules relating to the word spellings. These rules often have to deal with exceptions.
  - *Language Modeling:* This makes models like finite state machines using tools such as Markov chains for word ordering and even letter ordering in word. Probabilistic models are often made on real world datasets to realize them. They help in all further tasks. To capture the effect of sequencing, the basic entity considered is often a continuous string of 2-3 words called n-gram. Such models also help in the parsing procedure for syntax and even sentence level tasks like literature generation.
- *Syntactic Analysis:* Natural languages are infinite. It is not possible to store each and every sentence along with its interpretation in limited memory size. The only way to handle them is to have principles which guide how to form longer and longer sentences. At the heart of this study is syntax. So, this basically deals with analysis of sentences, their structure and function. Some subtasks are:
  - *Part of Speech Tagging:* Parts of speech have been introduced in the previous section. Strategies here are of three types: rule based, statistical and transformational[38]. The rule based methods use grammatical restrictions while statistical methods use solutions like hidden markov model, maximum entropy markov model or conditional random field. Transformation based strategies continuously keep transforming their tagging until no more corrections can be made, they can be hybrid in nature. Owing to derivation rules disambiguation is very important for tagging.
  - *Structure Defining:* This basically tries to teach computer all the rules for sentence formation. Abstract solutions like sentence diagrams are used to represent acceptable sentences. As described in the previous section major structures can be enumerated while minor ones need exception handling. These set of rules are called "context free grammar"(CFG).
  - *Parsing:* A parser is a recognizer that produces associated structural analyses ([e.g.] parse trees) according to the grammar. In a parse tree top down or bottom up approaches are used to identify if a given sentence is derivable in the given grammar or not. During the process not only the sentence is verified but also its grammatical structure is identified which helps in case based reasoning further and syntax directed semantic analysis.
- *Semantic Analysis:* It deals with the study of meaning in sentences and beyond. It is difficult to completely separate the study of semantics from syntax. Thus the structure and the basic constituents play a major role in understanding meaning. *Principle of constitutionality* states that meaning of a sentence is a function of meaning of its basic components. Since the meaning of meaning and understanding are too meta considerations for NLP, often a representation and reasoning system is first proposed and according to its ability to represent real world observations its credibility and usefulness are assigned. Thus the subtasks are:
  - *Meaning Representation:* Before reasoning is pos-

sible there must be a system which can represent meaning. The representation has to be verifiable, unambiguous, canonical, inference-enabling and expressive[?]. Usually logic languages are used to prescribe the knowledge base for a representation. Some common languages are first order logic, descriptive logic and modal logic. Semantic networks have been used for two decades now in knowledge representation since the inception of semantic web research. To deal with real world objects and abstractions ontologies and hierarchies are also used.

- *Interpretation Assignment:* It is possible to have different interpretations for a same concept. Formal concept analysis and logic uses domain dependent interpretation functions to map objects from meaning representation to real life objects. These have been found in almost all the logic systems. This is often considered a part of *Pragmatics* where out of the multiple possible meanings one of them is assigned considering context and intent.
- *Reasoning:* Reasoning is used on knowledge bases to derive new sentences and concepts. Reasoning systems must be sound as well as complete, i.e. they can entail all and only those inferences which are coherent with the knowledge base. Logic languages automatically satisfy this requirement. In the past decade there have been projects which deal with making knowledge representations which can offer common sense reasoning (Sys, ConceptNet, WordNet among others) but with limited success. Refer[38][39] for a more detailed analysis of knowledge representation and reasoning.

Some Pragmatic tasks: Owing to the developments in NLP and growing knowledge economy it is used at multiple places. Some of its basic applications are listed below:

- *Sentiment Analysis*
- *Summarization and Textual Entailment*
- *Natural Language Generation*
- *Natural Language Understanding*
- *Question Answering*
- *Machine Translation*
- *conversational bot*

H. *RETHINKING CULTURE AND CONTEXT*

Context is "a frame that surrounds the event and provides resources for its appropriate interpretation."[40]. It is a heavily studied topic in linguistics and sociology. This definition aligns with our society of mind ideas and provides us ground to interpret context in a manner such that it can be broken down and understood by our agents. Context can be of multiple types, a proper ontology is proposed later. It is easy to deal with context when it doesn't require knowledge specific to a culture or a society because then it can either be characterized by the particular case we are talking about or through the structure of the language itself. When context is cultural then it broadly affects the following[41]:

- *Pattern of thoughts*- Common ways of thinking, where "thought" Is meant to include both factual beliefs as well as values and emotional attitudes.
- *Pattern of behaviour*- Common ways of behaving; including behavior from ways of talking to ways of ploughing the fields.
- *Pattern of material manufacture*- Common ways of producing and using material objects; including objects from pens to houses.
- *Trace on nature*- Durable traces in Nature which are the result of the activity of members of the community toward the natural environment.

Understanding of the possible effects cultural context can have is crucial for a machine mind because often the interpretation of natural language presumes this understanding. This varies from understanding of affects differently across cultures to specific actions in a culture. In sociology, cultures are often classified as high context or low context depending on the background knowledge required to understand it. This knowledge might help in understanding of language because for all practical purposes an architecture may or may not use certain agencies while trying to understand a cultural action or construct. This behavior is analogous to humans when they deal with people from different cultures. India is typically considered a high context culture while USA a low context one and it is natural to assume different strategies to understand their contexts too.

III. PROPOSED WORK

This section is an attempt to extend the machine-mind framework. We start by describing frames and we lay down a questionnaire using which frames can be analyzed in our machine mind. Then we provide a theory for the emotions in the system. This theory would help the system in both realizing the emotions through natural language and be able to use these emotions as a resource for optimizing its own tasks. This is followed by an analysis of the components of the Z*-Numbers which was not done previously in the proposing papers. Some prominent questions regarding them have been answered here. After this, we describe some general purpose agencies. These agencies along with other specific agencies then are used to extract Z* numbers from natural language as well as consolidate multiple Z*-numbers. The section concludes with some cues for knowledge representation.

A. *CHARACTERIZING FRAMES*

Minsky[3], gives a vivid example of how frames can be used in language by using the scene from a party invitation. Over the years, frame knowledge representation has become a prominent area of research; it has been implemented multiple times and has inspired the present Object-oriented representation. Usually frames consist of slots which are variables that need to be filled with values. Values itself might be other frames, concept etc. Thus, a concept often has a hierarchy of frames that deals with its different variables. A full analysis of this hierarchy and how the components interact lies in the domain and knowledge representation and

is beyond the scope of this work. We aim to characterize how a frame works and responds to new data.
As seen above, frames can broadly be categorized as either syntactic, semantic, narrative or thematic for the context of language. The last two types are somewhat less general compared to the first two and its possible to describe many more such frame types. However, we stick to this classification.

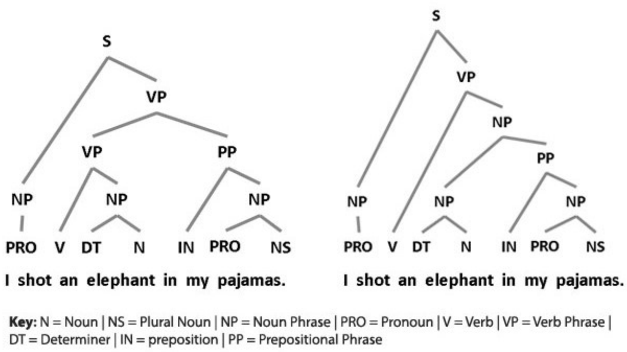


Fig. 6. A Sample Parse tree[?]

For an intuition on how frames can be useful, think of the sentence parse structures as shown in Fig. 6.[42]. A sentence tree is like a syntactic frame. The frame fitting is guided by constraint-satisfaction on word, phrase-level by using parts of speech and standard sentence structures. Once the frame is decided, one can do a syntax guided semantic analysis. The memories associated with the frame aid in the process. Thus, identifying the frame is half work done. Similarly, semantic frames represent concept and the associated context. For instance, a birthday party frame would specify the attire and acceptable actions in a birthday party. This way frames form the basis of the bottom layers of Minsky's model. Procedures and declarations via frames come in instinctive and learned flavors.

To fully characterize a frame structure, table VI suggests the questions that need to be answered about the machine mind and its frames (assumed to be following a society of mind architecture). A full understanding requires to specify the knowledge representation. We give simple intuitive answers in accordance to our machine mind and in the following sections assume that a full characterization has been laid down. Moreover, its only in the context of text understanding. The questioning framework would however go long in guiding the frame research.
(A) The structure of a frame greatly depends on the kind of frame it is. For linguistic purposes syntactic frame's function has to resemble that of a parse tree as shown in Fig. 6. To do so, usually the frames are prototypical grammatical formulas, which represent sequence and relation between the structures of the sentence like noun phrase, verb phrase, prepositional, etc. This naturally causes the part of speech tagging too. Semantic frames aim to deliver action-centered meanings of words, qualifiers and relations involving participants, instruments, trajectories and strategies, goals, consequences and side-effects. Their structure involves tables that have slots resembling propositions and values similar to objects. The relation in every table is linked by operators like is-a, for, of etc. Intuitively the slot can be anything like the abstraction written above and values be another kinds of frames or constants linked by a specific relation. They are like tabular representations of semantic networks. Thematic frames relate to topics, activities, portraits, setting, outstanding problems and strategies commonly connected with a topic. Thus the structure is very similar to semantic frames. However, in their case the values are more often the frames because concepts are often linked to other sub-concepts. The attributes of a theme, might be a theme itself. This calls for a hierarchical representation for a theme. The lattices used in formal concept analysis are analogous to the structure above. It is even possible to create intent and extent frames, thus generalizing or specifying a concept. Narrative frames differ from thematic frames in terms of attributes and specially spatio-temporality of the description. They constitute 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. The spatio-temporality can be taken care of using location/time attributes. For other slots, simple constants or thematic/semantic frames would suffice. Some attributes can directly be represented using syntactic frames because the parts of speech are functional in nature. This is true for the last two frames. Thus, besides the syntactic frames there is not much to say about the structures of the frames which are mostly tables, differentiated by the attributes in slots and the types of values. A detailed analysis would also talk about the nature of propositional relations, domain, range of values, etc.
(A1) For the syntactical frames its possible that due to environmental conditioning, some of the linguistic formulas modify themselves by merging two objects together or do a reordering. For, Instance the British English framework might receive an update to its formulas when subjected to African-American English where the use of negation and prepositions is pretty different. Thus changes are either sequence reordering or object granulation in formulas. Semantic frames are mainly propositional granules, thus a change in procedure or declaration would naturally update them. As far as complexity is concerned the structure would not change for the latter three frames for it is a simple slot-value kind of relationship. The hierarchical or network like relationship might become less or more complex though. For instance, the frame for a story in Ramayana which involves Arjuna, would clearly become less complex in its description if Arjuna, a character value becomes so. This is like pruning lower level leaves of a tree.
(A2) There is nothing more to mention for syntactic frames except that they are restricted to updation subject to the local grammatical rules. The rules are local because they have to be learned in a new environment and can't be rigid. An intuitive rule set is however provided, in line with the universal grammar theory of Chomsky[43]. For

TABLE VI

QUESTIONS THAT NEED TO BE ASKED TO A FRAME SYSTEM

Level of Description	*Question*
Structural	Q A. What would the frame look like? Q A1. (*Updation*) How is frame complexity (i.e. slot number, associations, description, etc.,) affected by previous frame associations? Q A2. (*Restriction*) What part of the structure can change and to what extent? Q A3. (*Modality*) How do frames differ with modality of concept representation?
Algorithmic	Q B. *(Choice)* How is the frame chosen? Q B1. (*Information*) How is an information-gain based approach used for choosing an appropriate frame and what are frame stimulants? Q B2. (*Memory*) How is the long-term memory involved? Q B3. (*Learning*) How does the system learn from bad choices? Q B4. (*Modality*) Can inter-modular frames be used and how? Q B5. (*Decision*) How is the exploration-exploitation trade-off between selecting new frame and improving on the current frame dealt with? Q C. (*Fitting*) How does frame-fitting occur? Q C1. (*Information*) How is an information-gain based approach used for fitting and what are the information cues? Q C2. (*Memory*) How is the long-term memory involved? Q C3. (*Membership*) Is the fitting process crisp or fuzzy? Q D. (*Creation*) How and when is a new frame created? Q D1. (*Improvement*) How are the old frames improved with time? Q D2. (*Pruning*) How and when are old frames pruned?
Implementational	Q E. (*Storage*) How are the frames stored in the memory? Q E1. (*Activation*) How does frame activation occur? Q E2. (*K-Lines*) How are frames attached to the K-lines? Q E3. (*Hierarchy*) How is the hierarchy implemented (i.e. data structures, reasoning, etc.,)? Q F. (*Machine-Mind*) Which parts of the machine mind control frame association and associated tasks? Q F1. (*Resources*) How are cognitive resources used; thrashing dealt with? Q F2. (*Restrictions*) Is design of other resources restricted due to the design choice of the frame architecture? Q G. (*Temporal*) Is the system sequential or parallel? Q G1. (*Order*) How are frame actions (on) and associations (with) co-ordinated? Q H. (*Automaticity*) Which parts of the framework are rule-based and which are statistical? Q H1. (*Ensemble*) How are mixed strategies implemented? Q H2. (*Interaction*) How do frames interact with each other?

semantic frames, obviously incoherent relations should bot be derivable. That is the representation should be logically sound and complete. This would restrict many changes like specifying multiple ages for a person, etc (for certain slots fuzzy relations are possible). This logical coherence would also pass on to thematic and narrative frames. Its not always required of them to have crisp relationships among variables, though. Following the natural subjectivity they would often take fuzzy values. Thus any part of the structure can change as long as logical coherence is maintained in the given knowledge base.

(A3) It is absurd to imagine that inter-modal conversion is possible without loss. Different modalities share different levels of description and occupy different resources of the mind while perception occurs. It is not in the scope of this paper to describe all different modality frames but we give some intuition. Frames are prototypical representatives of concepts. The representation can be visual, auditory, olfactory, etc. For instance a frame corresponding to a room would have a basic structure of the room and relative positions of various objects specified. It might also be paired with changing frames, constituting a video like concept. These ideas formed the basis of pixel representation later. Such a storage is of value because it removes the need to process the similar environment in the senses again. Only changes can be marked into the frame. Also they help in spatial-understanding and movement. This is because in their absence we can't predict any geography without sensing it again. For other modalities too the description is similar but instead of a pixel like representation with markings, the variables and the parchment itself is very different. For text things remain as suggested above.

(B, B1) Choosing the right frame for representing the data in hand, reduces to become a task of the difference engine if the current state is dynamically the frame of choice and goal state is the data in hand. The differences can be measured as information gain or loss, by only keeping the matter of the text which fits the frame. For instance, while fitting a syntactic frame only a unique frame would ensure maximum information gain. In any other scenario there would be a grammatical contradiction. In other cases the working is more complex. Thus the full description of choice boils down to characterization of the difference engine. A difference engine needs to have a measure of information gain or reward and a strategy to optimize the convergence both in reward achievement and time taken to do so. Information metric is modality and type dependent. The algorithm however can be universal subject to the condition that operationalization is uniform across all cases. For eg. a statistical algorithm like multi-armed bandit problem.

(B2, B3) Memory is useful because mere rule-based or statistical working of a difference engine algorithm is bound to fail for large amounts of data. A procedural memory for the game or a declarative one for a query would aid by initiating the engine from a probably accurate state. Moreover, the choices and actions taken in the past can be logged in to support the choice. This way the chance of a bad choice is decreased in future again. Knowledge-base would also allow the system to learn more about the surrounding this way and increment itself.

(B4) It is possible to use inter-modality frames for a description when it is not naturally called for. For example, an image of a forest would say much more than a thematic frame around the concept. Association of a frame with other modality frames in the past is a direct memory cue to do so. It naturally happens in the human mind too. A continuous translation between modalities would thus save time and build better understanding of the concept itself. This suggests that with every frame there are attached cues to the frames of all modalities like star with the image of sun. Over time, pruning would occur and only the most optimal representations would survive in the long term memory. Hyperfiles[?] were built on a similar concept.

(B5) It is based on two factors, the information gain associated with each action and the past memory of such decisions. It is possible to define a statistical measure as used in reinforcement learning too here.

(C, C1, C2) Actually, if the choice was made properly then fitting is already done. This is because the measure of goodness of our choice is the better fit itself which is based on the information gain or reward maximization. Over that, it is possible that minor tweaks further increase the information gain. This is difficult to formalize and has to be learned over time. It is analogous to the human process of learning to do to integration, and using better algebraic modifications with time. Thus long-term memory is involved in both major and minor tweaks but with latter its importance is formidable.

(C3) The process is definitely crisp when we talk of syntactical frames. With semantic frames crispness is decided by the crispness of specific value-slots. For the latter two however, multiple memberships are possible across different fittings. This is because natural language descriptions are often not precise and have natural subjectivity. Thus a concept can be classified into multiple categories at once. For eg, fair is less good as well as less bad subject to the context. This also establishes that the membership is a function of memories, belief association, context and of course the information gain. The fuzzy reasoning would go a long way in high level reasoning, because it often helps to looks at a object from multiple directions while reasoning about it. Same goes for themes and narratives.

This leaves us with a very seductive representation which can be fuzzy as well as multi-modal. However, it makes it even more crucial to have optimal representation strategies which are computationally feasible and don't cause thrashing at any level of memory. We also conclude that characterization of the difference engine is of vital importance and would be extremely handy while dealing with choice and fitting actions. Now we talk about the updating the set of frames itself.

(D) A new frame should be created intuitively when none of the available frame fits a given data with information gain or rewards above a particular threshold. The nature of threshold is event-specific and would be affected by memory too. The natural question is that how does one start creating a new frame? It would be impractical to create a new frame from scratch. rather cues from the failed fittings are taken. For the case of syntactic fitting it is trivial because most elements would have been at least identified i.e. tagged while using different grammatical formulas. They can be verified to follow the existing rule database and depending on belief attribution the KB must also evolve. Now a new formula can be constructed using the POS tagging. For other frames it is not easy to describe the process. However the intuition remains the same, cues which have been verified along with changes and cues in KB must direct the construction of the new frame. For semantic purposes formulas are replaced by propositional relations and for the latter two by more complex relationships. For the latter two new frames means new ways to think about concepts. That is because they are more artifacts of representation than of mere facts. It would be wise to describe the latter two only when structure and other operations are synchronized with the knowledge representation.

(D1) When a difference engine fails to find a better match but better information gain is possible using minor tweaks, they must be done. This is improvement in the existing frame. It can be temporary or become permanent as the frame evolves over time and data. Again the nature of tweaks has to be learned over time and stored into memory.

(D2) Pruning is done naturally when a frame regularly fails to offer solutions and can not even be tweaked to improve results. It can be temporary for the phase while choice is being made, depending on the difference engine strategy (just like pruning a game tree). Implementation-level questions can't be answered to their entirety without KR&R. Especially the ones related to storage of the frames (E, E1, E2, E3) and temporality (G, G1). We give intuition about other questions.

(F, F1, F2) All the tasks related with frames basically form the *"Connection"* subtasks of text comprehension i.e. *frame generation/retrieval/manipulation* over the layers of syntax, semantics, theme and narration as well as *encoding/decoding* of frame systems into customized knowledge components for optimal operations along with *memory handling*. We have also seen that there are two kinds of frame associations: local and global. While the deducer is responsible for manipulation, retrieval and generation using sub-agencies of syntax and semantics, manager is responsible for encoding,decoding and memory handling. The long-term memory components besides the lexicon are all directly related to frames. The *question-answer library* is a set of procedural frame knowledge while the *concept* and *commonsense networks* are networks of semantic, narrative and thematic frames in general. System memory management constructs also include frame pointers and hence access the frame system. Thus, each of the agency deals with frames in one way or other. [5] Talks a bit more about frame handling in machine-mind. many cognitive resources are used but if we look carefully only the lower 4layers of the mind are ever used during association. During creation however, all the six layers might be in use. There needs to a resources balancing between frame tasks. Again manager is handy by offering job scheduling to frames actions and associations. The design of other resources especially the memory constructs is definitely bound by the design of the frames. That's why we have deferred both of them for now.

(H) All the tasks of primitive association have to be rule based but as soon as there is a trade-off situation statistical or memory-aided pruning is required. For the purpose of fitting itself information gain, a statistical means is required. For creation and updation tasks a hybrid strategy is followed.

(H1) Mixed strategies comprise *multiple associations, ensemble associations* and *contrastive associations*. multiple associations would arise when either information gain or memory directs to use multiple frames and all the possible interpretations are equally likely. Ensemble strategy would be used when a frame alone can't achieve desirable information gain. This has to be intuitive again. Contrastive associations would be made when the sole purpose of association is compassion between two interpretations. Thus this is not truly a mixed strategy. Besides these fuzzy association may also be carried out. It would be similar to multiple association except there would be a membership engine in between. Mixed strategies are expected to aid in the design and improvement process too by giving novel insights to the system. In fact,the processes of uniframing and transframing which unify and bridge frame concepts are sort of mixed strategies. Them being a type of learning justifies our assertion.

(H2) Basically frames interact with each other via frame pointers, which fill in the slots of some frames forming a tree like structure. Other than that, two different concepts interact at a common frame. Z*-numbers ideally must be the token of interaction and also capable of becoming frame modules, in that ability they become the means of communication. However ample work is required to make them capable of doing so. This forms a part of this study where we enable them a little more and our future endeavors too.

Conclusively the concepts of information gain, difference engines and KR&R are required before we can proceed towards a robust description of frames. The naive looking details would pave the way for future analysis of the frames and Z*-numbers themselves. In the following section however, we assume that the frame system has been completely characterized.

B. EMOTIONS IN THE MACHINE MIND

For describing the emotion theory for the machine mind as well as answer the questions suggested above for an artificially intelligent system one needs to specify various components of a memory. However, if we think naively then only the few later theories of emotion like appraisal, core affect, network and concept act theory contribute to the machine mind's version because only they have had rigorous experimentation and have neuro-scientific evidence. We defer this important task to future.

C. Z NUMBERS REVISITED*

The following commentary on the parameters is mainly with regard to natural language. Dealing with multi-modal data forms a prospective work in near future. One important observation here would be the use of English grammar which specifies the domain for usage, however it must be noted that once a successful framework can be laid it is simple to extend it to other languages. In fact the task reduces to knowing the language grammar and collecting knowledge about the associated culture much like learning a new language in humans. Regarding grammar it must be noted that it has been historically used in natural language processing by computer scientists. Unlike rigid rule based and ignorant statistical solutions, ours is a hybrid system for natural language understanding. The more cognitive aspects of the process would surface when we deal with the agent architectures. In a way our analysis is very universal because it deals with basic cognitive processes rather than language specific constructs.

*1) **Subject**:* It seems that anything can constitute the subject for a Z*-Number for in previous descriptions[6] it has been so. However, this is not the case and at least for the context of natural language its possible to restrict what can constitute a subject in a particular case. Following cases are possible (Since its not a pure grammatical analysis it is possible the list is actually not exhaustive):

- When an interrogative sentence is being used followed by the an answer or cues to it, then the question may form the subject and the answer would be the attribute.
- When some object is being described, then the attribute would be the grammatical attribute or adverbial and that object would form the subject.
- When some process is being described, then the adverbial would be the attribute and the process converted into a suitable form the subject. The suitable form might

be a question or the verb may have been converted into a noun phrase.

Table[?] shows some examples of these kind of sentences and their corresponding subject. One interesting point to notice here is that this inconsistency between mentalese and natural language, of a possible exhaustiveness of subjects, is because of the limited capability of a language against thought, partly due to its grammar.

*2) **Time**:* It was assumed in the previous works and yet not explicitly mentioned that the parameter of time has a two-fold significance. It can express the time in the sentence as well as the time of the thought. The time in the sentence talks about when what happened/existed in the sentence happened/existed. This is indicated using the tense in the verbs of sentence (main or auxiliary). The time in the sentence is thus a grammatical construct. On the other hand the time of the thought shares characteristics of a mental attribute. This includes two components subjective and objective. The objective component is merely the system time and is used for logging operations. The subjective component on the other hand talks about the relative or subjective interpretation of time, not just in the sense of past, present or future but even finer compartments. This aids in the process of reflection and self reflection just like objective time, because memories would exist over granules of time rather than fixed points. The reason for this is consolidation of memories over time, flushing the record of the original experience. Thus, subjective time keeps track of the seamless mental travel of thoughts across time. As a novel idea we can add the concept of rate of thought. This allows us to define operations like fast forward and rewind on thought. From a conscious self-analysis they seem pretty intuitive because always keep doing such things with our conscious thought or the voice within. The rate of thought combined with emotional intensity at the moment would also imprint the memory and affect its later retrieval. A more deliberated thought sequence is more likely to be retrieved in future as compared to a thought that came and went away in passing.

*3) **Context**:* In alignment with the above characterization we suggest the following hierarchy for context for all practical natural language tasks. The specific agents which deal with these tasks have been described later.

- *Lexical and Grammatical:* This kind of context is embedded in the language itself. Lexical hasn't been used in the traditional sense and refers to any relationship which can be mapped like a dictionary. Grammatical context refers to rules of the language, which might convey some meaning at time. For instance, a certain grammatical structure might be linked to an emotion or a culture and can serve as a pointer for the same. Eg- idioms and proverbs, some local language structure like dialect, domain specific jargon etc. With regard to extraction/consolidation there are two possibilities here, either this context simply creeps into the Z* number or it gets mapped to a certain interpretation while doing so. It is roughly equivalent to saying whether it was syntactic or semantic in nature. The processor needs to decide whether the structure of the sentence itself has a bearing on the interpretation or not. We'd defer this task to past experiences and try to frame the problem in hand, in one of the known models. In many cases when the mapping is available (semantic context) as in an idiom or a sarcasm embedded in the sentence, the simplest interpretation would be carried ahead.
- *Cultural:* It forms the knowledge about the concepts of a culture like the festivals, stories, history, customs, science etc. Culture specific linguistic features, part of lexical context point to the specific cultures. The cultural context is the most diverse context and is especially required in understanding human interactions. It also consists of the ideas and thoughts which a culture has developed in years of human transcendence accompanying growth. There can be sub-cultures within a culture. Hence, a basic hierarchy is required. Similarly temporal association is required too. To deal with this kind of context we define something called a cultural network[?] as shown in Fig[?].

- *Case-specific:* This form of context includes all case-specific constructs[?]. This often talks about the people, surroundings, objects involved in the conversation. Famous Proper nouns are often included in the cultural context than in the personal context. It can further be classified in the following way.

  - *Emotional:* Context associated to the emotions expressed in the matter. It is different than AG because the latter talks about the inferred emotions via interpretation. Often the phrases or as we call them Emotionally Functional Words (EFW) are responsible for such expression. This is much easier to deal with then AG, because by definition it is a context not something which must be derived using other resources. While talking about the setting in which the sentence has been used, it often just describes the emotions of characters involved.
  - *Relational:* It talks about the the people involved in the conversation and the corresponding relations between them. It serves as a clue to entering the cultural network.
  - *Symbolic:* All the conversations that occur before or after a particular conversation and influence other types of personal context. Practically this is the context which people refer to while making conversation because it is less abstract and can be easily termed as "prior knowledge" for a conversation.
  - *Situational:* Description of the premise and the present understanding of the situation where the people are right now. This is very crucial in the sense that when one has a frame for interpretation that guides it in a setting, situational understanding helps to realize the direction where one needs to go further.

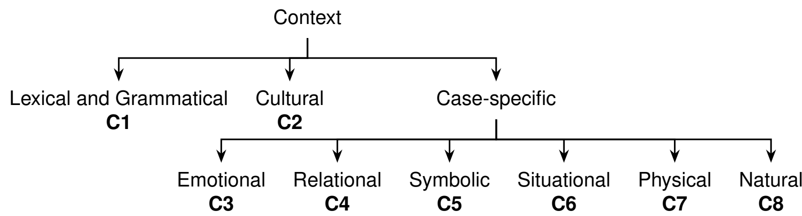


Fig. 7. Types of Context

- *Physical:* It describes the materials, objects and surroundings involved in the conversation.
- *Natural:* This broadly includes the other features of multi-modal communication like voice pitch, articulation, signs etc. These haven't been dealt in the present analysis.

Understanding these kind of context clues is simpler as compared to cultural context. Often this analysis helps in belief attribution to the agents involved in the act[6], which can be further passed on to the process of certainty analysis. Sometimes, memories about the agents/actions/environment can be cited to mine crucial context from memory. Just like cultural context, ideas not immediately visible in the sentence become a part of the context in Z* number here. This gives it an extra touch of intelligence and the framework transgresses mere knowledge based reasoning. The context classification has been summarized in Fig. 7.

*4) **Certainty**:*

*5) **Affect Group**:* The analysis of this component would require us to provide our theory of emotion for the machine mind.

D. GENERAL PURPOSE AGENTS

Some general purpose agents which aid all the language tasks have been defined below.

*1) **Parts of Speech Tagger**:* This agency collaborates with the frame-association manager to tag the parts of speech in the natural language. This is necessary because this demarcation is used in a case based analysis of different components as well as syntactic and semantic breakdown of complex problems into easier smaller ones. Clearly, syntax-surface frames have to be used in this task because they specify the grammar formulas. These formulas as shown in Fig. 6., naturally do the POS tagging. However, it is required that a certain level of tagging has been already done so that the frames may proceed via information gain. In NLP usually three kinds of strategies are used: rule-based, statistical and transformational. Since our work here involves frames as well as agencies we'd be using a mixture of rule based and lexical(statistical) approaches. The policy however would be like transformational, because we aim to do simultaneous error-correction after every step. The statistical approach here calls for the presence of long-term memory where we would have a kind of probability-distribution for every word against its possible part of speech. The distributions for other n-grams would exist too so as to consider sequentiality. However, it is crucial to notice here that usually while reading one is not consciously tagging words, rather the whole meaning is grasped as one. It is a result of years of training and finding shortcuts to the actual process of tagging. Exact procedure is not known to this date. Such level of expertise would require evolution of thinking and modeling strategies and we aim to provide them once our dynamic and plastic knowledge representation is in place. For now, some of the steps in the process are mentioned below:

1) Using the Memory the system recognizes the words/n-grams which can be tagged most confidently. (A HMM or CRF might be in use here, but details are not necessary at the moment.)
2) Using information gain a syntax frame is selected which maximizes the reward on the tagging. IF not found, creation of new frame or minor manipulation occurs according to the previous section.
3) The untagged words are tagged using the frame.
4) If there is any conflict due to frame tags then we either have to manipulate the frame itself or change the tagging. Confidence and belief attribution have to be used for this decision. Memory of such conflict handling would also aid.
5) After repeating the 4^{th} step multiple times, if any word is left untagged then lexical search is required to update frame/memory for such tagging.
6) Repeat 4^{th} and 5^{th} steps until all words have been tagged.
7) Log entries are cleared and memory updated.

Resources used:

*2) **Conjunction Analyzer**:*

*3) **Special Modulation Detector**:* It deals with linguistic modulations distinct from the general use, like sarcasm and figures of speech. It would be wrong to say that this is an agent in its own, rather it is a polyneme or a hierarchy of polynemes which have been associated with this task of language analysis in the past. This agent can be described further using more micro-tasks. That forms a future work. However, even in the present description it is self-sufficient because most humans learn about sarcasm and other modulation via experiencing them, forming a memory and then recalling it again in future. In the machine mind this forms a specialized network of K-lines which have some frame terminals and some cases of episodic memory.

Resources used:

*4) **Information gain analyzer**:* This is a more general and powerful super-agency which is used by various other agencies in language interpretation and usual frame association tasks. As described in the previous sections it is of vital importance but a complete analysis is possible only when multi-modal knowledge representation is fully characterized.
Resources used:

*5) **Memory Controller**:* This is a sub agency of the manager itself and is being stated again just for the sake of brevity. While dealing with any form of language the memory is not only used in interpretation but also gets updated. The act of linking stimulus with the memories is also a very important task. Informally, it controls and directs the memory traffic. Some of its sub agencies are:

- *Connection updater:* This updates the connections inside the memory based on the inferences you make from the data through the natural language. It may cause reinforcement, weakening, elimination, etc, for these connections.
- *Connection builder:* New connections often need to be made when a new kind of data is observed.

Resources Used:

*6) **Tense Analyzer**:* The analysis of tense is crucial for finding subjective time in natural language. This would be done using the lexical memory. As illustrated in the NLP survey item-process strategy must be used for long-term memory storage. Besides the modification rules of grammar, we also use frames. Frames and rules work until full constraint satisfaction is done. Finally, for difficult structures which are unlikely to arise for this case which is heavily formalized, we can use past associations or a reference to external lexicon and information sources like web can be made. Mentioning such rules and characterizing the memory association is a future task.
Resources Used:

*7) **Q&A Agency**:* This agency is responsible for finding answers to questions. In case no answer is found it returns that signal. A text document might be given or the machine might be prompted to "think" about the answer. It is not possible to specify this agency full without KR&R and we only mention it because it is used by many other agencies.
Resources Used:

*8) **Abstraction controller**:* The primary work of this agency is to control the level of abstraction in other agencies. The different levels can be symbolic, syntactic parsing, phrases, sentence, clauses etc. The control keeps shifting with more and more new data becoming available. This is a very powerful general agency which works for many other tasks in the architecture. The reason it is used here is because of multiple layers of possible interpretation in language.
Resources Used:

E. EXTRACTION: PROCESS AND AGENTS

Z*-numbers have been shown to successfully model the endogenous thought arousal[7], basic reasoning using language[6] and do most of the language tasks within the mind. However, it is of vital importance to actually extract Z*-Numbers from the natural language, because only then can they be utilized in these further tasks. This section provides a set of agents which carry this task while using the frames, knowledge representation and other cognitive resources of the machine-mind. Some of them even though not fully characterized, have been assumed to be functionally complete for the sake of brevity. Missing links would be filled in future work.

Frames are the primary source of interaction with the external data we receive in the form of natural language. Frames help in classifying the problem in hand. Also using frames, correct activation of memory structures is possible which is required to aid the different components with proper resources. That is why we have described the frame association task in the previous sections.
There are 6 components of the Z*-Number and hence 6 different sub-agencies for extraction below. Each agency follows a similar style of description where different components of the language have been broken and used for extraction tasks. Each of them have further smaller agents that deal with elementary language functions and form a nearly exhaustive representation of the same. Some of them use other agencies too which are otherwise general purpose. There is some grammar usage here and there considering the machine needs to know it innately, before it can learn its own grammar. It helps it in breaking down the extraction task into multiple subtasks before a case based algorithm can be laid out. The framework described below is a hybrid of knowledge based and learning models for natural language mining. These agencies work in a constraint satisfying parallel fashion because all of them are interdependent.
Some subtasks of the analysis are subject extraction (SE), context extraction (CE), time extraction (TE), emotion extraction (EE), attribute extraction (AE), certainty analysis (CA). Compounds of a clause and what can you learn with them (Refer Table[?] and section[?]):

1) *Subject:* Indicates the participant. Excites SE, CE (for memory retrieval).
2) *Predicator:* Describes the process. Excites SE (verbals like gerunds), CE (for memory retrieval), TE (for tense association).
3) *Adverbial:* Describes the setting. Excites SE, CE, TE (for rate of action), EE, AE, CA.
4) *Direct/Indirect object:* Further indicates about the participants. They can be used as subject, depending on the emphasis in the sentence. Excite SE, CE (for memory retrieval).
5) *Attribute:* Describes the subject/object in the process. Excites SE, CE, EE (When emotional component is present), AE, CA, TE(verbal usage as noun).

Our analysis of different types of sentences will proceed as follows:

1) *Simple:* They have a single independent clause, so the above clause analysis is followed. Mostly, there would be formation of just one Z* number.
2) *Compound:* They have multiple independent clauses,

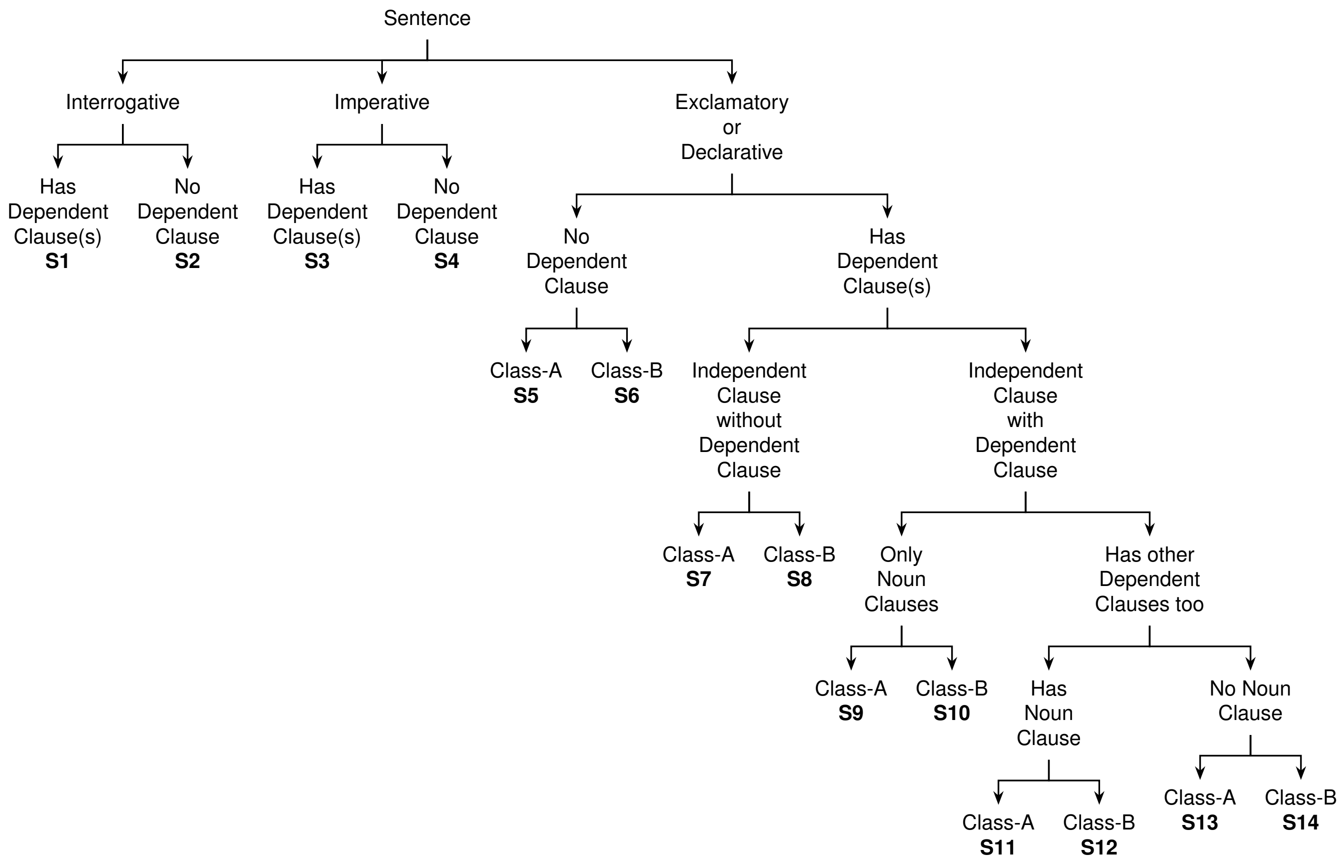


Fig. 8. Hierarchy of Sentences for Processing

along with coordinators. Two things need to be analyzed here:

- The nature of connector. A separate "conjunction analyzer" exists for it, as described above.
- The clauses. Two or more Z* numbers are formed from the clause analysis of the independent clauses. The input from the connector engine also goes into the analysis to indicate the presence of any subordination or coordination.

3) *Complex:* They have an independent and at least one full dependent clause. Use of a subordinator is thus mandatory. Subordinator defines the kind of relationship they share with the independent clause (analyzed in conjunction analyzer). They again have treatments of the following three types depending on the nature of the dependent clause:
   - Dependent adjective clause excites AE, EE, CE, CA.
   - Dependent adverbial clause excites AE, EE, CE, CA, TE (for rate of action).
   - Dependent noun clause excites SE (When they form the subject), CE (for memory retrieval).

Working with dependent clauses is a different subtask altogether. It would be described in the final framework or through the run-through because it is more algorithmic in nature.

4) *Compound-complex:* They have multiple independent and dependent clauses. Use of coordinators and subordinator is thus mandatory. This forms the most suitable part for analysis because it is the most general type of sentence. If the architecture can handle this general case all the above cases can be handled too.

We also need to analyze the process for different functional types of sentences. Before doing so we need to remind ourselves that Z*-Numbers were created to deal with objective and subjective components of the sentence. So lets look at the subjective elements once:

1) *Time* is subjective because human interpretations don't talk about objective periods or points in time. The subjectivity is in fact due to linguistic constructs and general precision of time perception.
2) *Certainty of attribution* is the original subjectivity Zadeh talked of[15]. It talks about the uncertainty associated in human speech which s difficult to express through mere probabilistic measures or statistical constructs; the very need for CWW.
3) *Emotions* form the fabric of subjectivity in human mind. They affect resource usage, job scheduling, memory actions, etc. These give the language subjectivity and aid in understanding action tendencies and motives behind language.
4) *Context* forms the subjective frame of analysis. It pro-

vides different interpretations and might be subjective itself because its encoded in a natural language.

When any one of them is completely lacking its not possible to completely extract Z*-Number. We propose that for such cases a partial extraction should do. How we deal with that in further computation is a future work. Bearing in mind these facts:

1) *Declarative:* The components of subjectivity vary according to the kind of declarative sentence we have (Refer section[?]). The different possible cases are:
   a) *Subject/Verb (Intransitive):* They don't use any object or subject attribute and simply describe a an action taking place. Thus subjectivity in attribution is possible only if the adverbial is being used to describe the setting. In case it is absent only partial extraction is possible.
   b) *Subject/Verb/Complement (Copula):* Since something is being said about the subject it is possible to have subjectivity owing to context and belief about it. The subjectivity in the setting might also be present but that is not the primary subjectivity. In some cases if the latter is equally important then multiple Z*-Numbers may be formed.
   Transitive verb forms are the only forms which can be used in passive construction. Thus, extra care is required.
   c) *Subject/Verb/Direct Object (Mono-transitive):* They have only one object, a direct object. Since there is no attribute the subjectivity of attribution can only come through adverbial.
   d) *Subject/Verb/Indirect Object/Direct Object (Di-transitive):* They too don't have an attribute thus only adverbial might contribute to the subjectivity of attribution.
   e) *Subject/Verb/Direct Object/Object Complement (Complex-transitive):* Both the object complement and adverbial can contribute to the subjectivity of attribution. However, former is the main contributor usually. Its still possible to have more than one Z* numbers.

   In all the above cases it is always possible to introduce subjectivity in the process itself. It would account for our beliefs about the participants and the process. However, such subjectivity doesn't care for the attribute, only certainty analysis is done. This again is partial extraction. We call the intransitive, mono-transitive or di-transitive verbs as 'class-A' and copula or complex-transitive verbs as 'class-B'.
2) *Interrogative:* This doesn't usually have an attribute or an associated certainty to talk of. If the text itself contains an answer then it is analyzed using frames. If it is a question directed to the machine-mind then it is answered thereon, the analysis for extraction is not required because either it was present in the question answer library or has been answered by seeking relevant context and reasoning over it. In the latter case all components of Z*-Number are ready and this problem can actually be deferred to answering questions using the machine mind which is possibly a future work. For the general case when answer extraction is not the burden we leave the components of A and B, the emotions elicited are stored in AG, the time subjectivity in T and the context in C. The act of questioning might have interjection though.
3) *Imperative:* This kind of sentence can't be used in text unless it is some sort of narration or is an instruction to the machine mind. If its an instruction there is no subjective factor of attribution or certainty. Other forms of subjectivity are dealt just like in interrogative sentences.
4) *Exclamatory:* This is an emotionally charged sentence, which may have all the components of subjectivity. There is usually a need to encompass interjection and exclamation in a specialized way while dealing with emotions. We'd deal with that in EE. Other components can be dealt just like declarative sentences.

Considering the above description about sentences, we propose a hierarchy in Fig. 8. for dealing with cases that may arise during extraction operations. As we can see in the diagram there are 14 possible cases, which we need to deal. Now we describe the extraction agencies.

*1) **Subject 'X' Extractor (SE)**:* It receives input from all the other parts of the sentences. Its task is to decide what is being described by looking at the sentence. As we have seen above multiple cases are possible. The way it deals with each of these cases is:

- In cases **S1, S2, S3** and **S4**, the whole sentence become the subject as discussed above. In cases **1** and **3** additional summarization is possible if context has a cue available.
- In cases **S5** and **S7**, the independent clause(s) become(s) the subject while removing the adverbial if it exists.
- In cases **S6** and **S8**, depending on what is being complimented subject or object become the subject. It is however possible that the whole sentence devoid of adverbial might become the subject depending on the emphasis in the sentence. These might be multiple Z*-numbers depending on how many independent clauses are present.
- In cases **S9, S11** and **S13**, the whole sentence forms the subject devoid of any adverbial if it exists. The noun clause might be treated later but is of no significance here. Any other dependent clause is not the part of the subject, its removed for brevity.
- In cases **S10, S12** and **S14**, the noun clause becomes the subject if it describes the object being complimented. Any other dependent clause is removed for the sake of brevity.

It is clear from the above discussion that frame association and POS tagging has to be precise in order for the agency to work properly because if the framing is wrong the case based reasoning would be wrong too. Some sub-agencies used by this agency are:

- *Emphasis Analyzer:* It decides how many Z* numbers need to be made from the sentence. Some factors to be considered are clause dependence structure, functional type, intent, context, emphasis, etc. In fact, context and subject are interdependent and follow a constraint satisfaction-like strategy. Using context signal and key-words which show emphasis some of the independent clause may/may not become a Z* number. Another important job is to decide whether the attribute or the adverbial is the correct description in this case. This again would require context and frame association. It is possible that different descriptions within the same clause form different Z*-Numbers because of different values in other parameters of subjectivity.
- *AE Messenger:* It signals the AE about what kind of subject is being used. Depending on that choice different subject different attributes might have to be used.
- *Q&A Agency:* This is a general agency and not a sub-agency which is used to detect the answers when a question forms the subject of a Z* number. The reason why that would happen is that interrogative sentence has been used. In case no answer is found a partial extraction is done which elicits the process of answering within this agency.
- *Frame Analyzer:* Besides the specified structures it is possible that a minor tweak might have to be made in the frame. This sub-agency controls that function.
- *Passive Construction Analyzer:* It deals with passive construction in the sentence. These areal possible only for transitive verbs. It is possible to identify a classification just like above but it has been skipped and can be used during actual implementation. This is present in every agency, but we'd not mention it again.
- *Clause Summarizer:* In many cases its possible to reduce the dependent clause into a smaller word group or even a logical derivative using the context and the memory elicited by the same. Sometimes for the purpose of attribution its possible to combine attributes from adverbial, attribute and dependent clause. This agency deals with that and passes the signal forward to AE.

*2) **Context 'C' Extractor (CE):*** As pointed above context is one of the must important components of the Z* number. It affects interpretation and is computed along with other parameters thus affecting even their extraction. It is of utmost importance in emphasis analysis, certainty analysis and actual meaning comprehension. A very vivid description of context factors was given above. Context has already been characterized in Fig. 7. This agency now needs to perform the following tasks:

- To identify the location of the context cues in the sentences using signals from different agencies and the use of frames.
- Elicit the correct set of memories and the cultural network as described below.
- To summarize and extract the context cues from the identified location to the Z*-Number.
- Perform a constraint like updation with other agencies while providing them cues for emphasis and other analysis.
- Use the memory controller agency to affect the state of knowledge using this episode.

Now all the above tasks are performed as follows for different types of contexts:

- In the case of **C1**, if the context is grammatical then it simply creeps in through the subject. This kind of context would be reflected when the machine mind contemplates over the frame being used. For example, the use of negation in African American English, is very endemic and would otherwise be grammatically incorrect. Now, if such a frame already exists in the system then the context would also be known. If it is not and a reference search is required to tweak the frame a bit then context would automatically achieved. When the context is lexical then similarly if it is present in the knowledge it would simply get extracted, if not then reference search would occur. Whenever a search occurs or a memory retrieval too, memory controller is used and memory is updated/reinforced. The location of this context is not fixed in the sentence and can be found anywhere. Usually if it is a special modulation like idiom then only partial extraction would occur because other components are simply futile and there exists a dictionary mapping. In case there is some kind of latent cultural reference that is passed to cultural context.

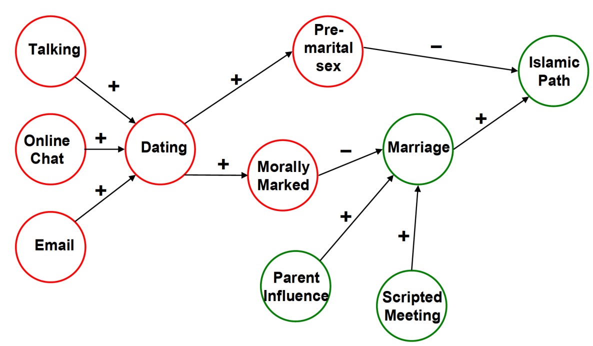


Fig. 9. Arab-American Culture model of romantic relationships [?]

- **C2**, is the most diverse and informative context. Practically, the specific strategy we suggest for cultural context here would actually be used in every other kind of context too but we'd by pass that now by saying memory retrieval. However, how is that information stored in the memory is very important for actually understanding the process. For the purpose of cultural context we suggest using something called a cultural network[44]. A culture network is a specialized semantic network. An example of the original cultural network is given in Fig. 9. It is very simple, the color of the

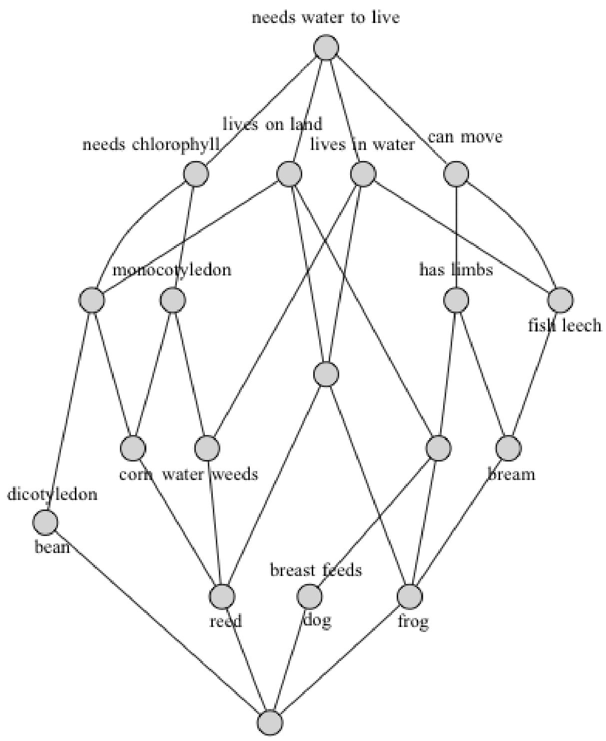


Fig. 10. Formal Concept Lattice[?]

vertex represents whether that is considered good or bad in light of a particular culture. The direction of edges represents the change former produces in latter. The positive sign increases the concept value or rather contribute to its characterization and/or certainty. This looks like a naive semantic network but this kind of graph can capture a lot of formal concepts. In fact in formal concept analysis similar graphs called lattices are used to represent formal concept. The backward direction represents the intent of a concept and the forward direction the extent. One direction leads to specificity and the other to generality. A formal concept lattice is shown in Fig[10][45]. We aim to make a new network with expressiveness of a culture network, reasoning capability of a formal lattice and with more subjective features. Some more features which we would like to add to the graph are:

1) *Time:* To add dynamicity, temporality and aid self evolution. Nodes/edges might be marked with subjective/objective time or different network trails could be initiated with a time trail.
2) *Emotion:* To understand the tendencies and possibly appraisal associated with the concept both from first person and third person perspective. The emotion associated with a change might also be considered.
3) *Fuzzy attribute:* It would be more useful to actually associate a fuzzy attribute with the node rather than label them green and red. This way it would be possible to convert this to CoW paradigm. It is even possible to associate it with multiple membership functions. Further, possibility and other certainty measures may be defined from this.
4) *Subjective Transformation:* It would again be beneficial to incorporate CoW in the transition of nodes. This is because there are levels of affect one concept might have on other. Moreover, its important to consider the accumulation of effects over time and node sequences.

On a closer analysis it seems like we are trying to identify all the other features of Z*-Numbers. In fact this network is just a temporary solution and in future work we want to aid the Z* numbers to be capable enough to form modular nodes of such semantic graphs. The result of such modifications would look something like Fig[?]. We don't formalize this anymore because it is futile to do so without complete knowledge representation and frame theory.

For the purpose of extraction whenever a cultural cue is obtained i.e. anything that represents gradual human normalization or typecasting, the associated network is accessed. For this kind of access to be computationally efficient it is necessary that in our knowledge representation we have an embedded ontology (some sort of hashing too maybe). Now all that is in the intent of the concept i.e. specializes it is taken as context (would define a proximity measure in future). It might also be possible to take the extent if the nodes are proximal. The associated signals for emotion, and attribute are sent to EE and CA. This way the cultural context is extracted.

- In the case of **C3**, the important keywords (EFWs) and associated memory are signaled to the EE. [46] presents analysis of such factors on which AG might depend.
- In the case of **C4**, the subject, object, noun phrases, etc., tell about the participants and their relationships. The memory associated with people forms important chunks of context like belief and attachment. Such cues also help the machine enter the culture network.
- In the case of **C5**, which is what is referred to as context in the natural language, memory cues have to be obtained. These cues can be obtained by observing keywords that refer to a previous state in time or a previous conversation. Again they can be present anywhere but would be most likely in the adverbial because this context is related to the setting and the participants. Clearly C4 would naturally retrieve past conversations. After the episode most recent conversation is saved into the memory, or rather its interpretation.
- In the case of **C6**, clearly adverbial the setting describer and the predicator the process itself would be the cue for the context. This context initiates the frames associated. All the above kinds of context are related to this one in the sense that memory associated with the situation would actually comprise of people, objects, conversation, interpretation, etc., only. It is analogous to interpretation in first order logic. After the episode

the most recent memory of situation is written in the memory.
- In the case of **C7**, objects and often the adverbial would be the location. This is totally dependent on the present and the past memory might not be completely true. However, the past frame allows fast analysis. This is analogous to the physical mapping of surroundings as described in visual frames by Marvin Minsky[3][4]. Memory gets updated and frames changed.
- **C8** is not dealt with right now. Analysis of natural context would in fact change the very upper ontology of context.

In [6] the analysis of context was based on a Z*-granule which was formed over the total certainty and affect of all attachment groups. It must be noted that while reading the text we have basically taken into account that information, because the memory associated with the event and reinforcement is actually related to the attachment figures' opinion and the belief attributed to each one of them. In future while we convert the Z*-granules into memory itself as was suggested in the culture context section, the memory recall and retrieval would actually be similar to averaging over attachment figures.

*3) **Time 'T' Extractor (TE):*** While we talk about extraction we don't really emphasize on finding the time of the thought but rather focus on time in language. To derive the normal notion of time we need to do a syntactic analysis of the sentence. More often than not, the verb's tense in the sentence would describe the time of the sentence. As mentioned above three kinds of subjects are usually possible. In all the three cases tense can be found using the general agency tense analyzer as described above. For the case of an interrogative being the subject the agency can target the sentence itself.

*4) **Attribute 'A' Extractor (AE):*** This agency receives input from adverbial and attribute components of the clause as well as the adverbial and adjective dependent clauses. Additionally, a signal from SE specifies the nature of the subject. According to the hierarchy explained above, following cases are possible for this agency:

- In cases **S1** and **S2**, it is possible to provide an attribute only when the answer is either in the text or known to the machine mind. Thus the Q&A Agency is nevertheless excited. If answer is not found partial extraction is done.
- In cases **S3** and **S4** attribution is not possible. If we get dead bent into finding one, then the adverbial or clause might provide one. However that is usually not the intent of the sentence, which is pretty objective in the commanding sense. Peripheral thoughts might be generated which include these less significant subjectivities. We'd deal with such Z*-Numbers in future.
- In cases **S5, S7** and **S9**, attribution is possible only when adverbial is present. The corresponding signal is received from SE.
- In cases **S6, S8** and **S10**, attribution is done using the compliment. Which compliment is to be used depends on the signal from SE.
- In cases **S11** and **S13**, certainly the dependent clause has to be the attribute because there is no other attribution.
- In cases **S12** and **S14**, the attribute might come from the the subordinate clause, attribute, adverbial or from a mixture of them. In case of a mixture clause summarizer would deal the case, in other cases the signal from emphasis analyzer would control attribution.

*5) **Certainty 'B' Analyzer (CA):*** In [6], *total certainty event* has been calculated by averaging over the certainty of all attachment figures while considering their associated beliefs. The working of this agency is similar but it works on the language level, because without cues to measuring the certainty of those attachment figures *total certainty event* can't be measured. That is because the level of representation of those beliefs is not with regard to the given sentence but rather in some abstract form in the memory, to access or relate to which we require linguistic cues. It is even possible that for other attachment figures such certainty is already present in the Z* granule or the memory but at least for the *self*, it has to be derived from the text. Following steps are followed to derive the linguistic cues:

- In cases **S1** and **S2**, attribution is possible only through the Q&A agency, thus when answer is provided internally memory operations take care of the associated beliefs by attachment figures.
- In cases **S3** and **S4**, attribution is not possible and thus not even the certainty. Partial extraction has to be done.
- In all the other cases the whenever attribution is not done certainty can't be provided too. In the case when mere declaration is present, there are no linguistic cues and the granular/memory operations have to be done to obtain the certainty of an event. More would be said when KR&R is dealt with.
- In the remaining normal cases when attribution is done, following inherent linguistic cues have to be searched for:
  - *Form of the auxiliary verb*, suggests a great deal about the inherent assumption on certainty. CoW analysis is required here.
  - *Modals*, like may, can, shall, must provide different sense of emphasis on a process and thus the certainty.
  - *Adverbial*, often contains the exact keywords for certainty like certainly, probably, etc.
  - *Structural:* Besides these cues sometimes a series of sentences via reasoning can be used to infer certainty.

More shall be said about the certainty once KR&R is in place and then we can merge psycho-linguistic factors with CoW theory.

*6) **Affect Group (Emotion) 'AG' Extractor (EE):*** Similar to certainty *Total affect event*, has to be measured through the averaging approach using attachment figure's belief in line with our theory of emotion. A deeper analysis would be given when the theory is actually in place. Following, sub-agencies are used for the linguistic cues:

- *Feature extractor:* The system needs to learn to extract relevant emotional features based on previous experiences. The relevance tagging machine can be based on information gain. The interesting problem however is that information captured through entropic measures can't be based on the data itself but the effect the data has on the system's previous connections. It seems like a more reflective task for the system. The formation of new nodes, new frames ad new connections is a rudimentary explanation of the same.
- *Interjection analyzer:* It works for symbolic as well as phrase/word based interjections. The symbolic interjections like ? and ! are easy to analyze. They provide information about the valence and possible modulations sometime. The other interjections in form of words are more direct and can be analyzed easily, like "eureka!!". A significant analysis of such words comes from formal training and storage in memory.
- *Declaration extractor:* It works like the interjection extractor. What is different is that the input it receives is in the terms of the residual information after conversion of the sentence forms into declaratives (exclamatory). Since all the symbolic information is being dealt by the interjection engine its work is restricted to ensuring no syntactic information is lost during conversion. When syntax is being changed some auxiliary parts of speech might get lost, which affect the affect expression.
- *Valence Controller:* Valence is one of the dimensions in our framework. To detect valence, we need to rely on the frames. But besides that, some features like the ones mentioned in [46] which came after some research on emotion detection are useful in optimizing this task and would also help.
- *Context analyzer:* Considering the presence of emotion on the boundary of social context and mind, it is very enticing to consider the possibility of interaction with culture network which forms the basis of context analysis. Some more details can be given here based on *folk theory of mind*.
- *Memory controller:* The kind of word and its previous usage. Also, the intensity of an emotion is expressed through different versions of an emotion e.g. Sad, mellow, grief etc. These concepts are stored in the memory in a WordNet affect kind of fashion and retrieval depends on which polynemes are instigated and finally revive which node. The emotion generated has some interaction with the culture network. The memory of previous such analysis i.e. a reflective memory is also used, in a sparse learner like mechanism it contributes to kind of the emotion. The complex emotions and expressions like sarcasm can be dealt with in this manner. One important task of the memory controller is to deal with the other sub-agencies and help them in making a contact with the long-term memory. The kind of words we are looking for here, are Emotionally functional words (EFWs):
  - Emotion Key words: Express an emotion or a hybrid of different emotions. A great lexicon can already be loaded from the databases like WordNet. e.g. Sad, mellow, grief etc.
  - Modifiers: Intensify or de-intensify a particular emotional feeling. These are adjectives, adverbs etc.
  - Metaphor words: These forms the analysis subject of special modulation detection framework. Their, mere presence can denote some emotions.
- *Multimodal engine:* The future work in the system development requires us to make these tokens self-sufficient to represent other modes of emotion expression like facial gestures, voice intonation, pitch etc. This module would be developed in future.
- *Syntax relation extractor:* Some sort of frame appraisal must be associated with extraction of emotions. These kinds of patterns need to be learned over time and be stored in memory in form of connections.

F. CONSOLIDATION: PROCESS AND AGENTS

The entire theory of consolidation requires us to specify the knowledge representation completely. A vivid and sufficient characterization would make the process a simple operation on KB. For example, if a cultural network is used than consolidation would reduce to cause-recovery and memory-reinforcement analysis. Circular relationships would be differentiated over historical usage of the particular memory. We defer this discussion to future.

G. CUES FOR KNOWLEDGE REPRESENTATION AND REASONING

Work for 2017-18/I whilst working in probabilistic machine learning.

IV. EXPERIMENTS AND RESULTS

Extraction needs to experimented with. Consolidation experiment has been done but would be added only after the theory is complete.

V. RELATED WORK AND DISCUSSION

REFERENCES

[1] A. M. Turing, I.Computing Machinery And Intelligence, Mind, vol. LIX, no. 236, pp. 433460, 1950.
[2] P. Langley, J. E. Laird, and S. Rogers, Cognitive architectures: Research issues and challenges, Cognitive Systems Research, vol. 10, no. 2, pp. 141160, 2009.
[3] M. L. Minsky, The society of mind. New York: Simon & Schuster Paperbacks, 2007.
[4] M. L. Minsky, The emotion machine: commensense thinking, artificial intelligence, and the future of the human mind. New York: Simon & Schuster, 2006.
[5] R. Banerjee and S. K. Pal, Text comprehension and the computational mind-agencies, Natural Computing, vol. 14, no. 4, pp. 603635, Jun. 2015.
[6] R. Banerjee and S. K. Pal, Z*-numbers: Augmented Z-numbers for machine-subjectivity representation, Information Sciences, vol. 323, pp. 143178, 2015.
[7] R. Banerjee and S. K. Pal, A computational model for the endogenous arousal of thoughts through Z*-numbers, Information Sciences, vol. 405, pp. 227258, 2017.

[8] S. C. Shapiro, Encyclopedia of artificial intelligence, 2nd ed. New York, NY: Wiley, 1992. pp. 5457

[9] R. V. Yampolskiy, Turing Test as a Defining Feature of AI-Completeness, Studies in Computational Intelligence Artificial Intelligence, Evolutionary Computing and Metaheuristics, pp. 317, 2013.

[10] T. Sethia, The lessons of ahimsa and anekanta for contemporary life: conference proceedings. Pomona: California State Polytechnic University, 2002.

[11] A. K. Chatterjee, A comprehensive history of Jainism. New Delhi: Munshiram Manoharlal Publishers, 2000.

[12] A. Jankowski, A. Skowron, and R. Swiniarski, Interactive Computations on Complex Granules, Rough Sets and Intelligent Systems Paradigms Lecture Notes in Computer Science, pp. 123134, 2014.

[13] P. Singh and M. L. Minsky, An Architecture for Cognitive Diversity, Visions of Mind, 2005.

[14] P. Singh and M. Minsky, An architecture for combining ways to think, IEMC 03 Proceedings. Managing Technologically Driven Organizations: The Human Side of Innovation and Change (IEEE Cat. No.03CH37502).

[15] L. A. Zadeh, A Note on Z-numbers, Information Sciences, vol. 181, no. 14, pp. 29232932, 2011.

[16] R. Banerjee and S. K. Pal, A Machine-mind Architecture and Z*-numbers for Real-world Comprehension, Pattern Recognition and Big Data, pp. 805842, Jul. 2016.

[17] P. Ekman, Basic Emotions, Handbook of Cognition and Emotion, pp. 4560, 2005.

[18] S. K. Pal and A. Dasgupta, Spectral fuzzy sets and soft thresholding, Information Sciences, vol. 65, no. 1-2, pp. 6597, Jan. 1992.

[19] E. Husserl, Husserls Text, Introduction to the Logical Investigations, pp. 1661, 1975.

[20] Predictably irrational: the hidden forces that shape our decisions, Choice Reviews Online, vol. 46, no. 02, Jan. 2008.

[21] Blindspot: hidden biases of good people, Choice Reviews Online, vol. 51, no. 10, 2014.

[22] J. Gottlieb, P.-Y. Oudeyer, M. Lopes, and A. Baranes, Information-seeking, curiosity, and attention: computational and neural mechanisms, Trends in Cognitive Sciences, vol. 17, no. 11, pp. 585593, 2013.

[23] G. Loewenstein, The psychology of curiosity: A review and reinterpretation., Psychological Bulletin, vol. 116, no. 1, pp. 7598, 1994.

[24] R. M. Stallman and G. J. Sussman, Forward reasoning and dependency-directed backtracking in a system for computer-aided circuit analysis, Artificial Intelligence, vol. 9, no. 2, pp. 135196, 1977.

[25] L. D. Erman, F. Hayes-Roth, V. R. Lesser, and D. R. Reddy, The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty, ACM Computing Surveys, vol. 12, no. 2, pp. 213253, Jan. 1980.

[26] J. de. Houwer and D. Hermans, Cognition and emotion: reviews of current research and theories. Hove: Psychology, 2013.

[27] M. J. Power and T. Dalgleish, Cognition and emotion: from order to disorder. East Sussex: Psychology Press, 2016.

[28] D. Marr, Vision: a computational investigation into the human representation and processing of visual information. CAmbridge: The Mit Press, 2010.

[29] Steunebrink, Bas R., Mehdi Dastani, and John-Jules Ch Meyer, ”The OCC model revisited.” Proc. of the 4th Workshop on Emotion and Computing. 2009.

[30] A. Ortony, G. L. Clore, and A. Collins, The cognitive structure of emotions. Cambridge: Cambridge University Press, 2015.

[31] J. A. Russell, Emotion, core affect, and psychological construction, Cognition & Emotion, vol. 23, no. 7, pp. 12591283, 2009.

[32] S. T. Scheuring and A. Agah, An Emotion Theory Approach to Artificial Emotion Systems for Robots and Intelligent Systems: Survey and Classification, Journal of Intelligent Systems, vol. 23, no. 3, Jan. 2014.

[33] Neviarouskaya, A., Prendinger, H. and Ishizuka, M. ”Narrowing the social gap among people involved in global dialog: Automatic emotion detection in blog posts”, ICWSM 2007 - International Conference on Weblogs and Social Media, 2007.

[34] E. C.-C. Kao, C.-C. Liu, T.-H. Yang, C.-T. Hsieh, and V.-W. Soo, Towards Text-based Emotion Detection A Survey and Possible Improvements, 2009 International Conference on Information Management and Engineering, 2009.

[35] Seol, Yong-Soo, Dong-Joo Kim, and Han-Woo Kim. ”Emotion recognition from text using knowledge-based ANN.” ITC-CSCC: International Technical Conference on Circuits Systems, Computers and Communications. 2008.

[36] M. H. Verspoor and K. Sauter, English sentence analysis: an introductory course. Amsterdam: Benjamins, 2007.

[37] K. S. Jones, Natural Language Processing: A Historical Review, Current Issues in Computational Linguistics: In Honour of Don Walker, pp. 316, 1994.

[38] Adhvaryu, Nidhi, and Prem Balani. ”Survey: Part-Of-Speech Tagging in NLP.” International Journal of Research in Advent Technology (E-ISSN: 2321-9637).

[39] J. F. Sowa, Knowledge representation: logical, philosophical, and computational foundations. Boston: Course Technology, 2012.

[40] R. J. Brachman and H. J. Levesque, Knowledge representation and reasoning. Amsterdam: Elsevier, 2009.

[41] A. Duranti, Rethinking context: language as an interactive phenomenon. Cambridge: Cambridge University Press, 2000.

[42] Allwood, Jens. On the role of cultural content and cultural context in language instruction. Vol. 60. University of Gothenburg, Department of Linguistics, 1990.

[43] R. Daniel, Parse Tree, New open access resource will support text mining and natural language processing. 2015.

[44] N. Chomsky, Aspects of the theory of syntax. Cambridge, MA: The MIT Press, 2015.

[45] W. Sieck, Cultural Network Analysis, Advances in Human Factors and Ergonomics Series Advances in Cross-Cultural Decision Making, pp. 260269, 2010.

[46] S. Willerton, Formal Concept Analysis. $https : //golem.ph.utexas.edu/category/2013/09/formal_concept_analysis.html$, 2013.

[47] Das, Dipankar, and Sivaji Bandyopadhyay. ”Identifying Emotional Expressions, Intensities and Sentence Level Emotion Tags Using a Supervised Framework.” PACLIC. Vol. 24. 2010.

[48] L. A. Zadeh, Fuzzy logic, neural networks, and soft computing, Communications of the ACM, vol. 37, no. 3, pp. 7784, Jan. 1994.

[49] Kotseruba, Iuliia et al. A Review of 40 Years of Cognitive Architecture Research: Focus on Perception, Attention, Learning and Applications. CoRR abs/1610.08602 (2016): n. pag.

[50] W. R. Sieck, L. J. Rasmussen, and P. Smart, Cultural Network Analysis, Network Science for Military Coalition Operations, pp. 237255.

[51] Tao, Jianhua. ”Context based emotion detection from text input.” Interspeech. 2004.

[52] B. Mesquita and M. Boiger, Emotions in Context: A Sociodynamic Model of Emotions, Emotion Review, vol. 6, no. 4, pp. 298302, Nov. 2014.

[53] E. C.-C. Kao, C.-C. Liu, T.-H. Yang, C.-T. Hsieh, and V.-W. Soo, Towards Text-based Emotion Detection A Survey and Possible Improvements, 2009 International Conference on Information Management and Engineering, 2009.

[54] T. Mitchell and E. Fredkin, Never-ending language learning, 2014 IEEE International Conference on Big Data (Big Data), 2014.