

## UNDERSTANDING THROUGH SENSE: COGNITIVE PARADIGM FOR COMPUTATIONAL ANALYSIS OF NATURAL LANGUAGE TEXTS

Olexii Voloshyn, Leonid Sviatogor, Vladislav Morgun

**Abstract:** *Within the Artificial Intelligence, this article considers the prerequisites of intelligent communication between two systems operating on knowledge in natural language form within a unified Linguistic world view. A new cognitive paradigm of Understanding through Sense has been formulated and will be evaluated through semantic analysis of natural language (NL) texts. The process of understanding has been interpreted as extracting Sense from the NL text by successive application of cognitive transformations to certain structures of knowledge, which are referred to as the “cognitive analysis units” (CAU). A text processing algorithm for the task “TEXT – SENSE” has been described in a general way.*

**Keywords:** *artificial intelligence, NL text processing, ontology, cognitive analysis units, text, sense, understanding.*

**ACM Classification Keywords:** *1.2.7 Natural Language Processing*

**DOI:** <https://doi.org/10.54521/ijita30-02-p02>

---

### Introduction

---

Establishing intelligent communication and mutual understanding between human and robotic systems using a natural language (NL) is a challenging task of Artificial Intelligence (AI). Rapid development of information technology has made this task of actual and resolvable. The humanization of society and creating a comfortable environment for human-machine communication is, in our opinion, an integral part of the global ecological process.

No decisive success in intelligent communication was achieved within the framework of separate modern sciences: neurobiology, mathematics, logic, linguistics, neurolinguistics, etc. Creation of a linguistic model “SENSE-TEXT” by I. Melchuk was one of the relevant attempts. [1] A failure to develop this model suggests that Sense is an extralinguistic issue. Presumably, neural networks are also an instrument rather than a theoretic platform to study the issue. It is a matter of course that human Intelligence is a unique and comprehensive phenomenon of social processes and humanitarian consciousness. A new scientific field, semiosociopsychology, is therefore expected to emerge [2]. However, the research efforts currently tend to be focused within cognitive science [3]. In view of this, it is necessary to mention the outstanding mathematician A. Turing, who was the first to formulate the problem of intelligent dialogue between two Intellects [5]. He was also the first to point out the fundamental distinction of Cognitive tasks from mathematical ones: human itself is a criterion of intelligent communication between people.

So what are the research resources of the sciences that aspire to study the human-level reasoning, intelligent activity, and communication? In this respect, several sources of knowledge are seen.

1. Natural language models represented in textbooks, usage rules for alphabetic characters and their sequences, as well as in oral speech and texts.
2. The input data for analysis are in most cases represented by NL texts, which record multiple historical implementations of social development, ideas, everyday situations, etc. These are typically literary works that contain knowledge and ideas that are important for society, have a theme, and are presented in a sensible, coherent form with a variety of styles. Such texts are considered weakly structured, as opposed to the rigidly structured texts of mathematical or statistical (tables) origin. In particular, structured texts are widely used in search and referral conversational systems. [4]
3. Instrumental resources are represented by Knowledge Representation Systems (KRS). These include semantic networks, ontologies, Linguistic world View (LWV), conceptual graphs, rule-based logic systems, frame-based structures, etc.

This study is based on the above resources. The general task comes down to the development of methods for cognitive analysis of NL texts. The purpose of the study is to: (a) determine the basic principles of intelligent communication; (b) determine the key concepts and objects of the task; (c) within the framework of the suggested paradigm “Understanding through Sense” offer and substantiate the working structures, which could serve as cognitive analysis units in the process of coherent transition from Text to Sense; (d) describe a generalized algorithm for recognition of the Sense of a free NL text.

Therefore, a solvable task is summarized as the development of a cognitive model “TEXT – SENSE”.

Cognitive analysis units (CAU) should serve as reference objects when describing the process for transformation of Knowledge from one form into another and deriving the Sense of an input message.

---

### **Principles of Intelligent Communication**

---

First of all, we should dwell on the prerequisites that distinguish intelligent communication from any other sort of communication. These are quite obvious.

It is evident that partners communicate within the common sphere of Knowledge accumulated in each society in the process of its cultural and historical development. Such global knowledge is called the “Conceptual sphere” [6].

For mutual and clear understanding, the partners communicate using the common language.

Every message carries its author’s Intentions, which is encoded in the NL text by words.

Sense is approaching the Intention. The Intention is a mental construct, which expresses a certain communicative Idea [7]. The Intention and Sense belong to the same Conceptual Sphere.

The Intention is formed within the personal Conceptual Sphere of the author Idea, while the Sense is synthesized within the mental Conceptual Sphere of

the partner. The two sets of knowledge should overlap at the level of the dialogue.

As a rule, the author of the text aims to transmit the information to the partner undisturbed – in the form of facts. However, sometimes the author may intend figurative meaning. In the latter case, additional intellectual efforts should be involved to restore the authors idea.

The key factor of contact efficiency, the prerequisite of intelligent communication is understanding a message.

---

### **Understanding the NL Text**

---

This category of Consciousness is a subject matter of psychology [8]. This article attempts at defining Understanding in terms of cognitive science as applied to computer analysis of NL texts.

First, we imply that understanding of information as a process means operations on Knowledge. Therefore, it should involve Knowledge Representation Systems (ontologies, conceptual networks or other KRS), which support specific Subject Areas (SA).

Let's suppose that the selected limited section of text displays a coherent fragment of reality. An appropriate Knowledge structure can be built for it. Let's call this structure an "input fragment".

Let's assume that there is a Knowledge Representation System (KRS) in the Knowledge Base of the Intellectual System (IS) and that the input fragment is presented in the same format as the KRS. Then you can set the "task of combining" of the input fragment and the KRS KB. For example, to search for a subgraph in a conceptual graph.

If it turns out that the input fragment of Knowledge coincides with the substructure of the KRS KB, that is, the "recognition" of the input fragment in the KB IS is recorded – then in this case it can be assumed that the input information has been "adopted" by the Intellectual System. The specified mental event of "adoption" of a fragment of Knowledge in a personal Knowledge Base

can be interpreted as an "understanding" of the input fragment by an Intelligent System.

Knowledge assumed as understood becomes an organic part of the Consciousness along with another knowledge.

Intelligent communication involves transforming subjective knowledge into objective one through understanding.

---

### **Understanding through SENSE**

---

While the concept "Sense" is widely used in everyday communication, this category of Consciousness is hard to objectify. In the context of the task for developing non-biological Artificial Intelligence, this concept should be formalized as an object of cognitive science. To achieve this, it should be made "computable" or "deducible".

***What is the role of Sense in human communication?*** During historical development of the language culture, a transition took place from the study of natural objects and relationships between them to the problems of interaction between Human and the environment (*Noosphere*). A branch of semiotics called pragmatics, which studies the relationship of the Sign to the User, has emerged. The pragmatics of the message suggests the extent to which information is useful for this Subject. Pragmatics fixes the individual and collective practical experience in human relations.

We classify Sense as a major category of pragmatics. Sense is unavoidable since it serves as a means to deal with huge amounts of data received by any person over the course of their life. Sense integrates and filters out the streams of external information, which emerge when observing the reality, and adjusts human behaviour to hazards and challenges in the actual environment. Sense is a destination of the cognitive analysis of reality, including the Linguistic World View. It consolidates information in the conceptual Knowledge System.

The Sense contains a cognitive interpretation of the essence of the current phenomenon — taking into account the surrounding context.

By abstracting itself from the artistic and stylistic “ornaments” in the text, the sense conveys to the consumer the main and essential content of the text, the main idea of the author of the message. Being short in form, it is both easy to remember and allows you to form an instant reaction. In this respect, Sense can be compared with an additional human sensory receptor, which provides a reflective response to an external linguistic stimulus.

We should distinguish between two types of Sense.

The “Generally accepted Sense” exists a priori in the Linguistic World View, since it accumulates the cultural and historical experience of Society.

It should be extracted from the memory.

In contrast, the "Constructive Sense" should be built in the mental Knowledge Base of the Subject using cognitive procedures of sequential transformation of Knowledge.

Each stage of these transformations is related to building of a certain structure – the “Unit of Knowledge”; the suggested system of Cognitive analysis units is described below.

**Hierarchy of Knowledge.** As previously stated, Knowledge representation systems may differ significantly from each other in their formal representation. The most developed are Ontologies. Conceptual ontologies are characterized by: (a) they are of a pyramidal type of semantic networks, (b) each conception forms a notion, except for the conceptions of the lowest level, (c) the structure of the network is formed by conceptual connections of the genus-species order, which determines a clear hierarchy of conceptual Knowledge. The latter property is extremely important since it allows to organize Knowledge according to the degree of their generalization. This makes it possible to conditionally divide the ontological pyramid network into horizontal "layers" so that the concepts of the upper layer give the set of concepts with which it is possible to "explain" or "define" the elements of the underlying layer. Consequently, we get a powerful platform for interpreting lower-level knowledge with the help of upper-level knowledge. In practice, this means that with the cognitive analysis

of the text, we get an ontological opportunity to generalize Knowledge. Such a procedure must lead to Meaning if it is implemented.

---

### **Prerequisites for construction of Cognitive Analysis Units**

---

A piece of reality is submitted to the input of the “thinking apparatus” in the form of sound signals, visual images, or a sequence of sentences from a NL text. In a localized and integral form, this piece goes through certain stages of cognitive analysis, until it receives the status of Sense. Knowledge that has acquired the format of a generally accepted or constructive Sense is considered “understood”, and the Subject may use it at his discretion - as “his own Knowledge”.

*Questions for understanding.* In real life, a person facing an unfamiliar situation usually produces questions: “who?”, “what happened?”, “where?”, “when?”, “why?”, etc. in order to understand it. Answers to these questions form a mental structure – the “image of the situation” (not always complete and clear). These questions cover only certain characteristics of the integral image (*Gestalt*). They are the answers on questions can therefore be referred to as the “*strata of Knowledge*” [9]. The strata of Knowledge in common are organizing a mental structure and fix this integrity mental object in subject`s memory as “easy to understand”, or “digested” one.

**Note.** We need the stratification of Knowledge in order to give unambiguous definitions and characteristics of the constructive *Cognitive Analysis Units* suggested here.

***Motivation of Knowledge.*** This cognitive operation is used in cases where a piece of reality can be related to its context. If the context of the fact includes causal relationships, which can be identified, such motivation can therefore act as evidence of information [10]. In cognitive science, motivation serves as a catalyst for understanding information.

---

### **The Units for Cognitive Analysis of NL Texts**

---

The following knowledge structures are proposed to define CAU.

### 1. NAME.

Strata of the Name: “who” / “what”.

Definition:

**NAME** =: *Parts of the sentence / Concept of the World Ontology.*

### 2. FACT.

Strata of the Fact: “Who”, “what does”, “object of an action”, “characteristics of the subject, object, action”.

Definition:

**FACT** =: *NAME.fct (Subject, Action, Object, Attributes).*

### 3. ATTRIBUTIVE FACT (IsA-FACT).

The canonical form of a Fact can be considered an Attributive Fact: “IsA-Fact”.

There are facts that can be distinguished from the rest as “canonical facts”.

They are often used in communication as directive definitions. For example: “The horse is white”. Where “is white” equals IsA.

Strata: “Who”, “what”, “what does he doing”, “which”, “whose”.

Definition:

**IsA-FACT** =: *(Name, Being smth, Attribute) =:*

*=: (Name, IsA, Attribute).*

### 4. EVENT.

An event is a cognitive analysis Unit that is semantically more complex than the Fact. To obtain an Event, it is necessary to provide the Fact with its “evidence”, i.e.: (a) the cause of its occurrence and (b) the consequence of its realization.

This motivation reflects the posterior (the Cause) and a posterior (the Consequence) characteristics of the cognitive process and forms a vector of understanding in the context space.

The Consequence in turn is of dual nature. If the action is not finished and continues to evolve, this naturally gives rise to a question “What is the Subject’s

Aim?” If the action is finished, the Aim is replaced with a Result. Therefore, the Event can be defined as follows:

***EVENT* ::= *NAME.evt (Cause, FACT, Consequence)* ::=  
 ::= *NAME.evt (Cause, FACT, Aim / Result)*.**

The Strata of the Event are formed by adding to the strata of the Fact new questions as follows: “Why”, “what for”, “what is a Result”.

## **5. SITUATION.**

We can assume that the author of a large message (work of fiction) builds its plot on the basis of events and situations. The user of information, in turn, must recognize these events and situations in order to build Sense based on them. It is evident that during this process of transition from Facts to Sense a strong reduction of data occurs.

Such intelligent consolidation of Facts (sentences of the text) and Events (“motivated facts”) is cognitive in nature, that is, it requires reflection and analysis. The Situation serves as an intermediate link in the analysis, which closes the set of Events into a single structure, giving it integrity to be varied with respect to the author’s Idea. This means that the complete discourse can be broken into one or many Situations, each of them matched with its Sense. The Sense of the complete discourse is synthesized on the basis of situational Sense.

Therefore, the Situation is a collective construction of a sequence of related Events. We propose to determine the structural representation of a Situation as an Frame construction [11]. The number of frame slots:  $N$  – is finite.

***SITUATION* ::= *NAME.sit (slot1, slot2, slot3, ... slotN)*.**

Situation slots are formed by questions as follows: “What happened?”, “who participated in the actions?”, “what are the parties’ objectives?”, “what are the parties’ instruments?”, “what are the results?”, “is it a final state?”.

This is a non-exhaustive list of questions. For example, when a criminal Situation took place is possible to add questions: “What are the artefacts?”,

“what is alibi?”, “who benefited?”, “who could be a criminal?”, “what I need to do in current situation?”, etc.

The Frame name is a “collective name” of the situation. The name of the situation has a significant cognitive load – it leads to an understanding of reality and finding the meaning of ongoing events.

The Situation is synthesized by applying a special cognitive operation – “Consolidation of Events”. This procedure is non-trivial in nature and currently requires further consideration.

As soon as the Situation is generated, we may proceed by formulating its Sense.

## **6. SENSE.**

*General definition.* Sense is the structure of Knowledge, which is included in the compendium of the written cultural heritage of mankind. It is used in communication to cognitively replace a detailed description of some situation.

This is a new category of cognitive semiotics with a clearly expressed pragmatic function.

Sense serves as an instrument for understanding the World, which transforms the events into personal experience (or Behaviour) of the Intellect.

To be a result of processing the data array, Sense – as it was showed above – can be expressed in two ways: (a) The “*Generally accepted Sense*” – as an aphorism of world culture and (b) The “*Constructive Sense*” – as a statement of NL, which preserves the syntactic, semantic and pragmatic characteristics ones.

### **Definition:**

(a) *The Generally accepted Sense (SENSE.general)* exists beforehand in the Human literature compendium of LWV (in NL form) as an aphorisms, proverbs and other forms of Knowledge.

For example:

*Actions speak louder than words.*

*The proof of the pudding is in the eating.*

*Don't count your chickens before they hatch.*

(b) *The Constructive Sense*. In contrast to the generally accepted sense, this type of knowledge is generated during the act of communication in the mind of the information consumer. In this case, the communicator relies on his mental Knowledge base and also uses a social cognitive resource. Constructive sense is also formed in the conceptual sphere of society.

**Definition:**

*SENSE.constr* is the structure of Knowledge, which is formed by an intellectual subject in his conceptual sphere and is contained in the mental Knowledge Base of the subject. Constructive sense uses concepts of high levels from the World Ontology.

Strata of knowledge: "what happened", "who / what", "where", "when", "why" (reason), "for what" (goal), "what does it mean" (estimation), etc. Obviously, the answers to these questions lead to an Understanding of the situation.

All these and other strata of Knowledge eventually boil down to one cardinal question: "What is the sense of what is happening?". Possible, at the next, higher level of cognitive analysis, a pragmatic question may arise: "What should I do?".

For example:

Thanks to a cognitive analysis of story "The 300 Spartans", a constructive complex Sense is received: "300 Spartans came out to defend Athens against the large Persian army. They knew that they were doomed to death, but they accepted the fight. They all died, but showed courage and detained the enemy. They accomplished a feat in the name of their homeland". A pragmatic question arises: "What should I do?" – There are two possible answers to this question in the Knowledge Base: "Follow their example.", "Retreat and wait for help."

---

**Conclusion**

---

The paradigm "**Understanding through Sense**" is implemented in an

algorithm, which is explicitly built upon the cognitive analysis units (CAU): **Text, Fact, Event, Situation, Sense**. The text contains as many Facts as there are sentences in it.

The original NL text contains N sentences, each of which is processed as a Fact. An array of N Facts is transformed into N Events using operation “Motivation”. An array of N Events is transformed into K Situations, where  $K \leq N$ . At this stage, the “Event generalization” operation is used. Sequence of N Situations is transformed into K situational Senses array. “Situational Senses generalization” operation is used on K situational Senses array. The result is a Generally accepted Sense of the text.

In the process of cognitive text analysis, sequences of simple CAU are converted into sequences of the high complexity level CAU.

CAU conversion scheme:

**NL Text** ► Seq. of Sentences (N) ► Seq. of Facts (N) ► Seq. of Events (N) ►  
Seq. of Situations (K) ► Seq. of Situational Senses (K) ►  
► **Generally accepted Sense (L)**

L is a quantity of Theses in Generally accepted Sense.

The CAU conversion operations listed above must be synthesized in the Knowledge base and are not considered in this paper.

---

## **Bibliography**

---

1. I. A. Melchuk. Language: from meaning to text / I. A. Melchuk. – Languages of the Slavic culture. 2012. – 76 p. -- (Studia philologica).
2. T. M. Dridze. Organization and methods of linguosociopsychological research of mass communication. — MSU, 1979. — P. 280.
3. M. Nesterova. Cognitive science: sources, challenges, possibilities: monography /Maria Nesterova. – Sumy: Universytetska Knyha, 2015. – 334 p.
4. T. A. Gavrilova, K.R. Chervinskaya. Extraction and structuring of knowledge for expert systems. – Radio i sviaz, 1992. – 200 p.

5. A. Turing. Can a machine think? – Fizmatgiz, 1960.
6. V. A. Maslova. Introduction to Cognitive Linguistics / V. A. Maslova. Flinta: Nauka, 2011. – 296 p.
7. Antoine Arnauld, Pierre Nicole. Logic or the Art of Thinking. – Kharkiv: Nauka, 1991. – 511 p.
8. I. V. Nastin. Psycholinguistics. – University of Psychology and Social Studies, 2007. – 180 p.
9. T. A. Gavrilova, A. V. Voinov. Visual conceptual analysis of knowledge and data for intelligent and instruction systems // 6th International Conference “Knowledge – Dialogue – Solution” (KDS – 97). Collection of studies in 2 volumes. V. 1. – Yalta, 1997.
10. Kr. Markov, Kr. Ivanova, I. Mitov. Theory of Infos. // Int. Book Series “Information Science and Computing” – Book No: 13. Intelligent Information and Engineering Systems, Sofia, 2009, pp. 9 – 16.
11. M. Minsky. A Framework for Representing Knowledge. – Mir, 1979.

---

#### **Authors` Information**

---

**O.F. Voloshyn** – Dr. Sci. (Tech.), Prof. Taras Shevchenko National University of Kyiv, 83000, Kyiv, Glushkova st., 4d,  
e-mail: [olvoloshyn@ukr.net](mailto:olvoloshyn@ukr.net)

**L.O.Sviatogor** – PhD (Tech.), Senior Researcher Taras Shevchenko National University of Kyiv, 83000, Kyiv, Glushkova st., 4d,  
e-mail: [lasvtgr@gmail.com](mailto:lasvtgr@gmail.com)

**V. V. Morgun** – PhD Student  
e-mail: [vladislavmorgun@gmail.com](mailto:vladislavmorgun@gmail.com)