Machine Understanding

From Wikipedia, the free encyclopedia

Introduction

Machine understanding, the term introduced by Z. Les and M. Les [15], [23], [29] [30] [31], denotes the process of understanding by the machine and is the first attempt to establish the scientific method to investigate the complexity of understanding problem. Machine understanding is referring to the new area of research carried out in a newly founded the St. Queen Jadwiga Research Institute of Understanding and initiated by Z. Les and M. Les [29] [30] [31], the aim of which, among others, is to investigate the possibility of building a machine with the ability to think and understand. Machine understanding defined in the context of both human understanding and existing systems that can be regarded as the simplest understanding systems is based on the results of philosophical investigations and assumptions of logical positivists. Logical positivists [34] adopted the verification principle, according to which every meaningful statement is either analytic or is capable of being verified by experiment and rejected many traditional problems of philosophy as meaningless. Machine Understanding is based on the development of the shape understanding system (SUS) [15], [23], [29] [30] [31] and on the assumption that the results of understanding by the machine (SUS) can be evaluated according to the rules applied for evaluation of human understanding. Machine understanding not only investigates the possibility of building a machine with the ability to think and understand but also makes it possible to study the selected aspects of understanding that can be approached using scientific methods and provides the suitable model of understanding. However, it is important to stress that machine understanding is based on the assumption that a machine can only to some extend approximate human understanding and there are some aspects of understanding that can not be approached by applying scientific methods. The most complex problem that is modeled within a machine understanding framework is understanding of visual objects (visual understanding). Machine understanding is defined in the context of both human understanding and existing systems that can be regarded as the simplest understanding systems.

Machine Understanding – Human Understanding

Machine understanding is defined in the context of human understanding. Understanding, the result of thinking, involves processes such as learning, problem solving, perception and reasoning, and requires abilities such as intelligence. Understanding (human understanding) is the topic of philosophical inquiries and only some problems related to human understanding are topics of research in the area of psychology, linguistics, cognitive science or artificial intelligence. Understanding appears as the result of the thinking process and can be the object of the scientific inquires according to Locke [32] who has no doubt that understanding can be studied like anything else. However, there are also some problems in understanding that can not be subjected to scientific methodology (empirical research). Human understanding was differently defined during the long period of philosophical inquiries. One view is that a perceived object and idea are a key to understand human understanding (Plato, Aristotle, Lock, Berkeley, Leibnitz or Kant (see e.g. [6], [32], [8], [1], [12], [11]). For Plato [37] understanding is grasping of ideas and the idea refers to particular things in the empirical world that are imperfect reflections of that idea, whereas for Aristotle understanding is connected with perception were ideas (concepts) are extracted from perceived data based on the abstraction and generalization. For Locke [32] understanding is grasping of the relations between ideas and for Kant [11] understanding begins by means of objects which affect our senses, produce representations, rouse our powers of understanding into activity (to compare, to connect, to separate) and to convert the raw material of our sensuous impressions into the knowledge of objects (ideas). For Husserl [34] meaning of the object is a key for understanding. He introduced distinction between natural and phenomenological modes of understanding. Natural understanding is regarded as the perception that constitutes the known reality whereas phenomenological understanding is considered as phenomenological reduction that is based on consciousness of any given object that discerns its meaning as an intentional object. For Frege [4], Wittgenstein [41] and Russell , language is a key for understanding and formal language and mathematical modelling were the important components of understanding. For analytic philosophy (logical positivism) [34] understanding is based on logical clarification of thoughts by analysis of the logical form of philosophical propositions and using formal logical methods to develop an empiricist account of knowledge. Logical positivists adopted the verification principle according to which every meaningful statement is either analytic or can be verified by experiment, and rejected many traditional problems of philosophy as meaningless. For hermeneutics philosophers (Schleiermacher [39], Gadamer [5], Heidegger [7]) interpretation of the text is a key to understanding. For philosophers such as Hobbes or Spinoza brain and its functioning is a key for understanding. Modern philosophers (logical behaviourism or functionalism) regarded the problem of understanding as the problem of mind functions and they and explains understanding by comparing the mind to a sophisticated computer system.

Machine Understanding – Simple Understanding System

Machine understanding is defined in the context of existing systems that can be regarded as the simplest understanding systems. Simple understanding systems are built in the areas of expert systems, image understanding, language understanding, or neural networks. Expert systems [9] are computer systems that emulate the decision-making ability of a human expert and are the first computer systems that solve problems that require understanding of the selected fragments of knowledge. The term image understanding [2], [35], [40] refers to a computational information processing approach to image interpretation and knowledge-based interpretations of visual scenes. Language understanding [38], [36], [33], [10], [3] is an area of research that deals with understanding of a text as the product of the linguistic activity of the mind. Neural Networks [42], a set of simple computational units (nodes, neurons) that are highly interconnected, which attempts to model the capabilities of the human brain, can be also regarded as the simplest understanding systems.

Machine Understanding

Machine understanding as a new area of research carried out in a newly founded the St. Queen Jadwiga of Understanding, aiming at investigating the possibility of building a machine with the ability to think and understand, was initiated by Z. Les and M. Les in the year 1999 [15], [23], [29] [30] [31]. The authors defined machine understanding in the context of both human understanding and existing systems, regarded as the simplest understanding systems. Machine understanding is based on the results of philosophical investigations and assumptions of the logical positivists. Further, machine understanding was established based on the framework of the shape understanding method and the development of the shape understanding system (SUS) [15], [23], [29] [30] [31]. The main assumption of machine understanding is that the results of understanding by the machine (SUS) can be evaluated according to the rules applied for evaluation of human understanding. The most complex problem that is modeled within this framework is understanding of visual objects (visual understanding).

Shape Understanding System (SUS)

A machine, in order to be able to understand, needs to imitate the way in which humans understand the world and language (text). Machine understanding has been established based on the framework of the shape understanding method and the development of the shape understanding system (SUS). SUS is designed to cope with problem of understanding of the visual object (the visual understanding) and meaning of the text. SUS as the machine that is designed to have an ability to think and understand learn both knowledge and skills. SUS operates based on knowledge of image processing, decision making and search strategies as well as knowledge of shape description and representation distributed among the specialized experts Les [14], [16], [24], [25], [26], [28], [27], [19], [20], [18], [21], [22]. The analysis is carried out by invoking the 'expert' that performs a suitable analysis employing a specific method. The expert performs the task given in a form of requirements by another expert based on the internal ability to use the knowledge of its domain of expertise as well as communicate the obtained results with other experts.

Machine Understanding – Problem Solving

Nearly all activities connected with understanding can be regarded as a problem solving. The problem is any statement formulated in any possibly way that requires finding the solution. Solving the problem is to find the solution that can be unique or can have more than one solution.

Machine understanding is based on the assumption that the result of understanding by a machine can be evaluated and compared to the result of human understanding regarded as the problem solving. If understanding is defined as the ability to solve a problem, then assuming that problems (tasks) are well defined the understanding (ability to understand) can be tested by testing whether these problems can be solved by the machine (SUS). In this context machine understanding can be viewed as problem solving, however the machine to be able to imitate the human understanding needs also the ability to explain how to solve a problem or to explain the causes, context, and consequences of given facts. For this reason, the most important part of evaluation of the machine's (SUS) ability to understand is to formulate the problems and to use these problems to test if the machine (SUS) is able to solve those problems. Examples of problems that are solved by the machine (SUS) are: the naming, solving the visual problems (perceptual problems, the visual analogy problems or the spatial problems), the problems of the signs interpretation, the problems of text interpretation or explanatory problems. One of the important problems solved within the framework of machine understanding is naming of the visual or the sensory objects. Naming can be regarded as solving of the problem of finding the meaning of the object. When the object is named, its meaning consists of all learned knowledge that is linked to the category to which the named object belongs. For example, understanding signs is to solve the problem of finding the meaning of signs or symbols and is based on the learned coding system. Similarly, understanding a text is to solve the problem of finding the meaning of the text

where the interpretation that is based on the learned script is utilized. Explanation can be also regarded as solving the problem of explaining known facts, perceived objects, solved tasks or interpreted texts.
In the context of defining machine understanding as the problem solving the most important part of evaluation of the machine's (SUS) is to formulate the problems and to use these problems to test if the machine (SUS) is able to solve those problems. There is the special class of problems used for testing the results of learning at school, called, in machine understanding, the text-task that is very useful to test understanding capabilities of machine (SUS). These problems (text-tasks) are often part of the texts (e.g. from handbooks or textbooks) the aim of which is to check the level of understanding of the material presented in the textbook as well as the level of understanding of learned knowledge to which the meaning of the text refers. However, in order to test if the task presented to SUS as the text is fully understood by SUS there is a need to formulate the special text-task in the form of questions, computing problems or explanatory problems. For example, in order to test the degree of understanding of the mineralogical dictionary text SUS can be asked the question "*what is the name of the mineral that is represented by chemical formula ZnO*"? or to explain "*why malachite is green*"?
Machine understanding regarded as the problem solving means to stress finding the solution to a problem. The problem can be given in the form of the inner question or the question given by a user. Understanding of the perceived object (visual understanding) usually starts with the inner question "*what is this?" or* "*what is the name of the perceived objects?* and refers to the meaning of the object represented by the name of the category to which the object belongs. When the object is identified as the real world object and named, the name of the object can be used as the reference to its meaning.
The important part of the problem solving is to find a suitable form of the problem representation. The visual representation, as one of the forms of the problem representation, can be used as the problem itself (e.g. naming), as the schematic representation of the problem (e.g. solving task with electric circuits), as the imagery transformation (e.g. solving task planning robot action) or as the explanatory process (e.g. explaining a solution).

Machine Understanding – Categories

Machine understanding refers to the categorical structure of learned knowledge. The machine (SUS) to be able to understand an object needs to convert the essential features of the object into the basic shape categories and to assign the object to one of the visual object categories. The visual object category is part of the perceptual knowledge on which all knowledge used in machine understanding is based. The knowledge of visual objects is the basic ingredient of all understanding processes (in machine understanding) and is also basis for the abstract categories. Following the result of philosophical inquires, thinkers such as Locke or Kant that all ideas formed in the mind are the result of the sensory impressions and that ideas, the result of the faculty of the mind called intuition, are formed based on the impression that comes from the abstraction of the sensory material, SUS intuition is related to the SUS perceptual visual field that is the fundamental basis of the basic abstract categories. The basic abstract category is represented by configuration of objects on the rectangular perceptual field. These visual representations that refer to the SUS intuition can be utilized during explanatory process and make it possible to found understanding on the strong intuitive basis. Following the way of scientific understanding, machine understanding is based on the basic abstract categories such as the set category, the element category, or the belonging category that are defined in the area of set theory based on adopted axioms [31]. However, the abstract categories, co-derived from the knowledge category can be also seen as ideas (e.g. ideal geometrical figures) that are imposed on the perceived material to form the meaningful sensory categories. Machine understanding is based on the assumption that there is a meaningful dependence among categories and that categories of the scientific knowledge have the same meaning in the different scientific domains. For example, the categories defined in mathematics have the same meaning in any other scientific discipline such physics or chemistry. For that reason the categories in physics are defined by application of the previously defined mathematical categories and basic physical categories. All interpretation as the meaningful understanding process will relate learned categories to the knowledge that was learned by SUS. The machine (SUS) to be able to understand needs to relate learned categories to the basic abstract categories. The basic abstract (intuitive) categories are the set theory (STA) categories such as the set category or the member category. The basic abstract categories are applied during abstraction in the problem solving when the perceived object is assigned (transformed) into the visual general abstract categories such as the circle category or the rectangle category whereas the shape categories are basis for the intuitive grasping of the sense of perceived objects. The abstract categories that represent the knowledge that comes from the direct perception of objects (intuition) are derived as the essential property of perceived objects and other categories are derived from the visual object category. The essential features of perceived objects are the shape categories (shape classes) that are derived based on the characteristic attributes of visual objects.

Categories of visual objects are established based on the assumption that a visual object exists and can be perceived by the accessible technical tools [29]. Categories of visual objects supply knowledge about the visual aspects of the world. The categories derived from the category of visual objects are called the visual categories whereas the categories derived from the category of the body of knowledge are called the knowledge categories. The notation of categories is based on a categorical chain. The categorical chain is a series of categories derived from the categories of visual objects or the categories of the body of knowledge showing the hierarchical dependence of knowledge. The category at the first level of the categorical chain of the categories of visual objects is called the perceptual category of a visual object. The category at the second level of the categorical chain is called the structural category of the visual object. The ontological category begins from the third level of the categorical chain. The perceptual categories and structural categories are associated with the visual appearance of an object and are represented by visual knowledge.

Machine understanding refers to different ontological categories of objects: a visual object, a real world object, a sign, a sensory object, or a text object [30], [31]. Ontological visual categories have hierarchical structure and at the bottom of each categorical chain is the prototype category. The prototype is defined during learning process at the level for which the training exemplars are available. The prototype is represented by all visual representatives of a specific category and it is assumed that learned visual knowledge is covering a visual domain prototype. The visual domain prototype refers to visual knowledge that makes it possible to recognize all visual representatives of the prototype. For example, the visual domain prototype for the category of capital letters “T” consists of all fonts and handwritten characters of the category of capital letters “T”, whereas the visual domain prototype for the category of the font aerial capital letter “T” consists only of the aerial font of the capital letter “T”.

The sensory object category is a special category derived from the category of visual objects. The sensory object is usually the visual object that is named based on the complex sensory data rather than on visual features of the object. The complex sensory data is a set of measurements that refer to the attributes of the category to which the object is assigned e.g. the category of mineral objects. The text category is referring to any form of the text and is divided into four different specific categories: the text-query category, the text-task category, the dictionary-text category, and the long-text category [30]. An object from the text category is interpreted based on the category of the body of knowledge. The category of the body of knowledge (knowledge category) is divided into the category of theology, the category of philosophy, the category of science (scientific knowledge) or the category of common sense knowledge. The category of scientific knowledge is divided into the category of physical sciences, the category of biological sciences, the category of medicine, the category of engineering or the category of social sciences.

Machine Understanding – Visual Understanding

The most complex problem that is approached within machine understanding framework is to understand a visual object (visual understanding). Visual understanding is based on the processing of a perceptual data or processing of the internal visual objects connected with imaginary processes (imagination). Understanding that does not involve processing of a visual object is called a non-visual understanding or simply understanding. Although understanding a text involves processing visual objects (letters) during reading and text pre-processing, the text understanding (without images) is regarded as the non-visual understanding.

The visual reasoning usually assigns the name from one of the visual object categories to the perceived object. The visual object, after naming, is interpreted based on knowledge of ontological visual categories and knowledge of the knowledge scheme. The visual reasoning is part of the higher level visual reasoning used in the naming process where all learned non-visual knowledge that is connected with the category to which the name is assigned to the object is accessible. Higher level understanding processes involve the reasoning that is based on the previously learned non-visual knowledge.

The first stage of visual understanding involves perceptual reasoning that consists of perceptual categorical reasoning and visual reasoning. The visual reasoning consists of the assigned reasoning that assigns the perceived object to one of the shape categories. In contrast to existing approaches in AI, where usually reasoning is independent of the acquired data needed in reasoning process, assigned reasoning consists of the consecutive stages of reasoning where at each stage of reasoning the specific data are acquired based on the results of reasoning at previous stages.

The perceptual categorical reasoning, the first stage of the perceptual reasoning, is related to the SUS perceptual visual field where an object is assigned to one of the perceptual and structural categories. The perceptual category reflects perceptual properties of the object, determines the visual reasoning process and is divided into a silhouette, a line-drawing, a colour object, or a shaded object. The method of assigning of the object to one of the perceptual categories, based on the histogram, depends on a number of peaks in histogram. In a similar way the object is assigned to one of the structural categories. The structural category refers to the complexity of the visual representations of an object and is divided into the element category, the pattern category, the picture category or the animation category

The visual reasoning consists of the assigned reasoning that assigns the perceived object to one of the shape categories and is based on the shape understanding method . A member of the shape category is called an archetype. The archetype is an ideal realization of the shape (visual object) in the two-dimensional Euclidean space. The exemplar is one of the regions of a binary image and is a binary realization of an archetype in the discrete space. Although it was assumed that a visual object is represented by a binary image it is not the cause of a serious limitation to the presented method. The visual object that consists of parts of different colours is assigned into one of the colour classes and during processing stages these parts are interpreted as the new visual objects. The assigned reasoning is the most important part of the perceptual reasoning. The assigned reasoning consists of the consecutive stages of reasoning where a perceived object is at first transformed into a set of critical points and next into the symbolic name. In order to fulfil the required task of acquiring the data and processing it in order to obtain a set of descriptors, a processing method is used. The processing method applies an image transformation in order to transform the data into one of the data types. The descriptor transformation is applied to find a set of descriptors used to assign the perceived object to one of the possible classes. The assigned reasoning involves transformation of the description of an examined object when passing stages. At each stage of the reasoning the following operations are performed: the processing transformation transforms the set of critical points, the descriptor transformation computes descriptors, an examined object s is assigned to one of the possible classes. During reasoning process, a perceived object is first transformed into a set of critical points and next into the symbolic name. The symbolic name is extracted from a symbolic description. The symbolic description is an intermediate form that has many additional specific data about the perceived object (phantom). The symbolic description is used to reason about the specific categories to which the object can belong. For example, the object O1  is transformed into a symbolic description in the form of the following string: "[A3][[|L3|AE|]|S79||B100,99,99||A60,61,60||G248||@2691|]]{[|L3|O|]|S52||B58,100,57||A29,30,120||G76||@395|]}{[|L3|O|]|S52||B57,100,58||A30,29,120||G76||@396|]}{[|L3|O|]|S53||B100,58,57||A29,120,30||G76||@417|]}".

Next, the symbolic description is transformed into the symbolic name given as the string A3_L3_AE_L3_O_L3_O_L3_O.

Understanding of an object is performed at two levels, the intermediate level and ontological level. At the intermediate level of understanding the object is described in terms of the shape classes. The description of the object at intermediate level refers to the symbolic name. For example, for the object , the symbolic name (in SUS notation) "A3_L3_AE_L3_O_L3_O_L3_O" consists of two parts. The first part "A3" gives a general description of the class that means that the object is the acyclic object with three holes. The second part "L3_AE_L3_O_L3_O_L3_O" gives a specific description of the object. The final description of the object, at the intermediate level of understanding, is given in the form of the linguistic description: "acyclic object with three holes". At this level, the object is also described in terms of the structural archetype. At the ontological level, the object is assigned to one of the ontological categories during naming process. Naming not only attaches the name to the perceived object but also "connects" the object with all knowledge that is relevant to the name of the object. Many names from different categories can be attached to the same object and naming can be given at many different ontological levels. In order to assign the object to the specific ontological category, information included in a symbolic description is used to obtain the additional data needed in the reasoning process. For example, an object  can be interpreted as a symbol "eye of dragon" when additional relation "*all three holes are equal*" is established. In the case of the object O1  the size of holes is given in the string form as |S52|, |S52|, |S53, as the part of the symbolic description. The object O1 can be also interpreted as a mathematical object (solid pyramid) or as a real world object (a model of a pyramid).

Understanding of the Object from the Real World Category

A real world object is the three-dimensional (3D) object, whereas SUS perceives the object as the image – the 2D object. Images (pictures) are the basic visual perceptual categories of SUS - the objects of the visual perception. The image (2D-object of perception) can be a photograph of a real world object, a scene or a phenomena, a medical image obtained by application techniques such as X-ray Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), or an image created by a computer e.g. computer graphics, drawing or painting.

Understanding of the object from the real world ontological category requires usually interpreting it as a 3D object. SUS understands a real world object as the object extracted from an image in the SUS perceptual visual field. The real world object is usually extracted from the visual object (image) that is assigned to the perceptual picture category. The object from the perceptual picture category that is to be assigned to the ontological real world category consists of real world object on the background. The different backgrounds require applying the different segmentation methods to extract the object from the background. Knowing an object (name of object) that we are looking for makes the searching for the object and extracting it more easy task.

The real world object is always perceived as part of the environment. SUS always perceives the object as part of the scene, given as the image (picture). The object is perceived as the object on the background and as the part of another object. For example, fish in a river is part of the river, the river is part of the land, and the land is part of the planet Earth. From the picture category the different specific categories of pictures are derived. The picture (image) categories are usually representatives of the real world scenes comprising objects from the real world categories. For example, the image of the macro-landscape category comprises objects from the river category, or the road category. Understanding object such as the road sign (object from the road sign category), that is also the object from the real world category, is to extract this object from the perceived road-landscape scene transformed into the image, and attach the names to them during the naming process. The object (sign) is extracted from the road-landscape scene during the segmentation process. During the segmentation process the image is divided into two regions: the figure and the background and as the result of the segmentation the sign is extracted.

Understanding Sensory Object

When understanding of the object from the real world ontological category requires usually interpreting it as a 3D object, understanding of the sensory object refers to an object that has a complex conceptual representation. A sensory object is the object from the category of visual objects that is named based on a set of measurements that refer to attributes of the category to which the object is assigned. Understanding (naming) an object from the category of sensory objects is to classify the object to one of the categories of sensory objects. For example, a sensory object that belongs to the category of minerals is assigned to the mineral category based on the measurement of the characteristic minerals features. The aim of naming (recognition or classification) is to assign the examined object to one of the mineral categories based on a set of measurements and finding the mineral category for which the measures of attributes of the object are matched with the values of attributes of the mineral category.

Text Understanding

Machine understanding refers to different ontological categories of objects and one of the most important parts of understanding of the object is to understand an object from the text category (text understanding). The text category is divided into four different specific categories: the text-query category, the text-task category, the dictionary-text category and the long-text category. Text understanding is related to language understanding that deals with understanding of a text as the product of the linguistic activity of the mind, however text understanding is defined in the context of machine understanding and follows the way of understanding within the machine understanding framework. Understanding the text is to answer the questions concerning the given text, to understand the problem presented in the text by finding the solution and to give explanation to the solution, to give an example that explains the difficult parts of the text, to explain the main issues connected with the text, to make the abstract or short description of its contents or to translate the text into other languages.

The text is interpreted in terms of its meaning. The meaning for the dictionary-text is usually given by the interpretational script, whereas the meaning for the text-query and text-task is given by both the explanatory script and the basic-form. Understanding a text depends on the text category to which the text is assigned. Meaning of the text-task consists of the two different parts: meaning of the text that refers to the real world situation (phenomena) and meaning of the text in terms of the task that needs to be solved. The first part, the interpretational meaning is given by the description of the stereotypical situation in the form of the explanatory script, given at the different levels of description that reveals the different levels of the details. The second one, the basic meaning requires to transform the text into the basic-form and to identify the type of the solution by transforming it into the procedural-form. The interpretational meaning of the text-task is given by the explanatory script.

The interpretational meaning makes it possible to explain why some parts of the text-task are not relevant in solving of the text-task. Meaning of the text can refer to the visual object, to the sensory object or to some phenomena or event. For example, meaning of the dictionary test CUPRITE:

"*Cuprite A major ore of coppe, cuprite is named from the Latin cuprum meaning "coper". It can turn superficially dark grey on exposure to light. Cuprite typically has cubic crystals. In the variety called chalcotrichite or plush copper ore, the crystals are fibrous and found in loosely matted aggregates. Cuprite is a secondary mineral, formed by the oxidation of copper sulphide veins. Fine specimen come from Namibia, Australia, Russia, France and the USA."*

refers to the sensory object – mineral cuprite. The mineral that is a member of the category of 3D real world objects can be represented by mineral's name – cuprite, the chemical formula *Cu2O (coper oxide),* or by a photograph – its visual representation. Understanding of the dictionary text such as CUPRITE makes it possible to understand the text based on the learned explanatory script.

Understanding the dictionary-text is based on the previously learned interpretational script, whereas understanding the text-task is based on the explanatory script, the basic-form and the procedural-form. Understanding a long-text such as a text that belongs to the category of the set theory texts (STA texts) is based on the learned script. The STA text is the mathematical text that consists of the text units, such as contents, preface, introduction, chapter or subchapter, or title. The chapter of the STA text consists of STA statements, such as: the chapter title, the definition, the notation, the remark, the example, the theorem, the lemma, the corollary, the property, the proof, the problem, the solution to the problem. The mathematical text consists of the symbolic mathematical statements, the mathematical statements, the linguistic mathematical statements and the graph. Understanding of the STA text is based on the learned explanatory script. Understanding a text does not mean to understand all statements, however the deep understanding that is needed during the explanatory process or translation requires understanding not only each statement but also each word, symbol and mathematical expression. Deep understanding of the text requires depth knowledge of the grammar, semantics, syntax, idioms of the source language, as well as the culture of its speakers.

Explanation

Machine understanding is based on the assumption that the results of understanding by the machine (SUS) can be evaluated according to the rules applied for evaluation of human understanding. In general, machine understanding is defined as the ability to solve problems assuming that problems (tasks) are well defined. When solving problems used to test if the machine (SUS) is able to solve these problems can indicate the capability of machine to understand there is another very important aspect of understanding that is connected with ability to explain how to solve the problem or to explain the causes, context, and consequences of given facts (machine explanation). However, the most important aspect of machine explanation is the ability to explain how to solve a problem.

In a broad sense machine understanding can be regarded as problem solving and, in this context, explanation can be viewed as solving of the problem of explaining how to solve the problem or explaining known facts. Explanations is usually defined as a sets of statements constructed to describe the sets of facts which clarify the causes, context, and consequences of those facts, cannot be given in one unique way and are often represented by different media such as music, text and graphics. In machine understanding an explanation (explanatory process) is regarded in two main aspects (explanatory modes): the first one is focused on explaining of the solved problem to a user (out-explanatory text), whereas the second one deals with understanding the explanation (in-explanatory text) given to the machine (SUS). The result of the explanatory process is the explanatory text (in-explanatory text and out-explanatory text).

Explanation of finding a solution (problem solving) to the task (command-text-task) such as "*solve an algebraic equation"* is often given in the form of the complete step-by-step explanation for the solution. The procedural form of this command-text-task can be represented as the transformation of the symbolic expression w into the final solution. This procedural form can be used to construct the explanatory form as the complete step-by-step explanation for the solution. The procedural form can be seen as a sequence of descriptions of actions that justify and explain those actions. Understanding involves also understanding the explanatory text given by a user when SUS is solving the simple command text-task such as "*explain it"*. SUS, during the explanatory-testing session, asks the user to explain each step of the solution, for example, for the equation 2x2+4x-5=0 the following queries can be given: "*is this equation* ?", "*why*?", "*show me equality sign*", "*show me the right side of this equation* " and the answer that SUS receives needs to be also understood.

Visual explanation, that is part of machine understanding, is to explain a problem by using visual representation of the problem. One of the forms of visual explanation is explanation based on utilization of the perceptual analogy or schematization. For example, schematization that transforms the Earth into a circle can be used in the explanatory process where the planetary system is seen as a set of circles. SUS understands the complex visual category by referring to the basic shape categories such as the circle category or triangle category. In this context the Earth (image of Earth) can be represented by its schematic equivalent – the circle.

Machine Understanding – Knowledge Implementation

Machine understanding is strictly connected with learning of new knowledge. The machine (SUS) that is designed to have the ability to think and understand needs to learn both knowledge and skills. SUS operates in two main modes, learning and understanding mode and learning knowledge and skills, which supply material for thought that leads to understanding, is called knowledge implementation ([30]). SUS ability to understand depends on the effectiveness of learning process and learning of new knowledge depends on the SUS ability to understand. Knowledge implementation stresses dependence of the learning and understanding process and is concerned with learning both knowledge and skills. Learning of the new skills is to learn both the new methods of solving the problems and implementing of the new methods of processing and storing of the acquired knowledge. Acquiring of the new knowledge involves storing the new knowledge in the form of selected knowledge representations. Learning knowledge of the visual object is to learn the visual concept, the categorical chain and the knowledge scheme. Learning knowledge of the sensory object is to learn the sensory concept (model), the categorical chain and the sensory knowledge scheme. Learning knowledge of the text object is to learn the coding categories, the query-form, the basic-form, the procedural-form, the explanatory script and interpretational script.

Understanding in SUS is based on the learned knowledge acquired during the knowledge implementation process. Knowledge implementation is concerned with two main aspects of human learning: learning of the visual knowledge in the context of the categorical structure of the learned categories of visual objects and learning of the knowledge that is connected with understanding of the content of the text. During learning of the visual knowledge the generalization, the specification, the schematization and the visual abstraction is important part of the learning process. Understanding the visual objects from one of the ontological categories requires learning of the visual concepts of this category. The ontological category is given by its name and is represented by a set of visual objects called the visual representatives of the category. It is assumed that a set represents all visual aspects of the category. Visual knowledge of the category is learned as a visual concept represented as a set of symbolic names. During learning of the knowledge of visual objects, at first, the representative sample of objects from the category is selected, then for each object the symbolic name is obtained and finally the visual concept of this category, as a set of symbolic names, is learned. As a result of learning the visual concept is obtained.

Machine Understanding – Aesthetic Evaluation

Understanding objects from the different categories requires many different skills and knowledge that is acquired during the process of knowledge implementation. For example, understanding music as the object of the music category is different from understanding music as the object that belongs to the category derived from the art category. When understanding music as the object from the music category requires learning of knowledge of the interpretation of the music as the form of the sound that is produced by musicians and perceived by listeners, understanding music as the object from the art category requires the knowledge and skills to interpret the music in terms of the artistic aesthetic qualities. Similarly, other objects derived from the art category such as the category of pictures (painting) is understood as the object subjected to the method of aesthetic evaluation. Aesthetic evaluation is regarded as one of the forms of understanding. One of the first attempts to build the system of aesthetic evaluation, based on the image understanding approach by Z. Les, was presented in [13] and [17]. In this method trained neural networks were used to perform a simple task of aesthetic evaluation. However, based on these research and the results of the research in machine understanding there is no ground to believe that any computer system can have the ability to make aesthetic judgement for more complex tasks connected with aesthetic evaluation. Aesthetic evaluation is one of the abilities that are reserved only for spiritual being – the human being. It was the reason why the logical positivists rejected many traditional problems of philosophy such as the aesthetic evaluation as meaningless or the nonsensical problem that cannot be approached by applying scientific methods. However, a machine performing the simple aesthetic evaluation can help better understanding of the complex problems that are solved during aesthetic evaluation of the true work of art.

References

1. Berkeley G (1996) Principles of Human Knowledge and Three Dialogues. The World's Classics. Oxford University Press, Oxford
2. Binford T (1982) Survey of Model-Based Image Analysis Systems. *International Journal of Robotics Research* 1 (1):18-64
3. Dale R, Moisl, H., Somers H (2000) Handbook of Natural Language Processing.
4. Frege G (1967) Concept Script, a formal language of pure thought modelled upon that of arithmetic. In: J. Hv (ed) From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931. Harvard University Press, Cambridge, MA
5. Gadamer H (1976) Philosophical Hermeneutics (trans: D. L). University of California Press, Berkeley
6. Gilson E (1991) The Spirit of Mediaeval Philosophy. University of Notre Dame Press, London
7. Heidegger M (1996) Being and Time. State University of New York Press, Albany
8. Hume D (1975) Enquires Concerning Human Understanding and Concerning the Principles of Morals. Oxford University Press, Oxford
9. Jackson P (1998) Introduction To Expert Systems 3edn. Addison Wesley,
10. Jurafsky D, Martin JH (2008) Speech and Language Processing. 2 edn. Prentice Hall,
11. Kant I (1996) Critique of Pure Reason. The Everyman Library. Everyman, London
12. Leibnitz GW (1996) New Essay on Human Understanding. Cambridge Texts in the History of Philosophy. Cambridge University Press, Glasgow
13. Les Z An Aesthetic Evaluation Method Based on Image Understanding Approach. In: The First International Conference on Visual Information Systems VISUAL'96, Melbourne, 5-6 February 1996.
14. Les Z Shape Understanding Based on the Multi-Level Representation. In: Hamza MH (ed) The IASTED International Conference on Computer Graphics & Imaging, Palm Springs, 25-27 October 1999.
15. Les Z Shape Understanding System. In: The International Conference on Signal Processing Applications & Technology, Orlando, 1-4 November 1999.
16. Les Z (2002) Shape Understanding System: Understanding the thin object. *An International Journal Computers and Graphics* 26 (6):951-970
17. Les Z (2010) Image Understanding and Aesthetic Evaluation, vol (electronic version). QJF Press, Melbourne
18. Les Z, and Les, M. (2005) Shape Understanding System: Understanding of the Complex Object. The Journal of Electronic Imaging 14 (2):023015-023011-023015-023013
19. Les Z, and Les, M. (2005) Shape Understanding System: Visual Intelligence Tests. International Journal of Intelligent Systems 20 (6):799-826
20. Les Z, and Les, M. (2005) Understanding of a Concave Polygon Object in Shape Understanding System. Journal of Computer and Graphics 29 (3):365-378
21. Les Z, and Les, M. (2006) SUS: 3D Interpretation as a Part of the Visual Concept Formation. Machine Graphics and Vision 15 (2):141-175
22. Les Z, and Les, M. (2006) Understanding of the cyclic object as part of the visual understanding in shape understanding system. Journal of Computer and Graphics 30 (5)
23. Les Z, Les M System of Experts to Perform an Epistemological Task of Shape Understanding. In: The IASTED International Conference on Software Engineering, Las Vegas, 6-9 November 2000.

24. Les Z, Les M (2002) Shape Understanding System: The Visual Reasoning Process. International Journal of Pattern Recognition and Artificial Intelligence 17 (4):663-683
25. Les Z, Les M (2003) Shape Understanding system: Understanding of the Convex Objects. The Journal of Electronic Imaging 12 (2):327-341
26. Les Z, Les M (2003) Understanding as an Interpretation in shape Understanding System. Journal of Experimental and Theoretical Artificial Intelligence 15 (4):433-459
27. Les Z, Les M Shape understanding system: Understanding of the Complex Thin Object. In: IASTED Conference Computer Graphics and Imaging, Kauai, Hawaii (accepted), 2004.
28. Les Z, Les M (2004) Understanding in Shape Understanding System. International Journal of Pattern Recognition and Artificial Intelligence 18 (4):1-19
29. Les Z, Les M (2008) Shape Understanding System. The first steps toward the visual thinking machines., vol 86. Studies in Computational Intelligence. Springer-Verlag, Berlin
30. Les Z, Les M (2013) Shape Understanding System - Knowledge Implementation and Learning, vol 425. Studies in Computational Intelligence. Springer-Verlag, Berlin
31. Les Z, Les M (2015) Shape Understanding System: Machine Understanding and Human Understanding, vol 588. Studies in Computational Intelligence. Springer-Verlag, Berlin
32. Locke J (1961) An Essay Concerning Human Understanding. Dent, London
33. Manning C, and Schütze, H.; (1999) Foundations of Statistical Natural Language Processing. The MIT Press, Cambridge, MA
34. Martinich AP (2001) A Companion to Analytic Philosophy. In. Blackwell, Oxford,
35. Matsuyama T, and Hwang, V. (1990) SIGMA: a Knowledge-based Aerial Image Understanding System. Plenum Press, New York
36. Mitkov R (2002) Oxford Handbook of Computational Liguistics. Oxford University Press, Oxford
37. Plato (1993) Republic (trans: Waterfield R). Oxford University Press, Oxford
38. Radford A (1988) Transformational Grammar: An Introduction. Cambridge University Press, Cambridge
39. Schleiermacher F (1998) Hermeneutics and Criticism. Cambridge Texts in the History of Philosophy. Cambridge University Press, Cambridge
40. Ullman S, Richards W (1989) Image Understanding. Ablex Publishing Corporation, Norwood
41. Wittgenstein L (1961) Tractatus Logico-Philosophicus (trans: Pears D, McGuinness BF). Routledge & Kegan Paul, London
42. Zurada J (1992) Introduction to Artificial Neural Systems. West Publishing Company, St. Paul, Minnesota