

ENHANCEMENT OF COGNITIVE ABILITIES OF AN AGENT-ROBOT ON THE BASIS OF IMAGE RECOGNITION AND SOUND PERCEPTION

Wladimir Stalski

Technical University (St. Petersburg), Russia.

ABSTRACT

The objective of this work is to enhance the cognitive abilities of an agent robot (ARb), a model of a human. This article considers a new approach to ARb simulation, consisting of the utilization of a direct analogy between the functions of specific organs of the human brain/head and the "brain/head" of a model, between the life of a human and the "life" of a model. The model of a Homo Sapiens is constructed as an intellectual agent integrated with a humanoid robot. The task set is to construct an ARb that must be able to read and write, to understand its situation in the world, the meaning of its actions, and the semantics of the text it processes.

An ARb achieves the high cognitive abilities mentioned above because, like a human, from the moment of its "birth" it learns a natural language, first and foremost such as the names of objects and phenomena it sees and recognises. It is proposed that an ARb will be taught, including language teaching, in a group of both fellow robots and humans, under the tutelage of a teacher.

The growth of the cognitive abilities of an ARb is also achieved thanks to the evolution of a "population of reproducing agents" (our article [1]). For "reproduction", it is proposed that we should separate the "private life" of an ARb from its operating functions ("service"), i.e. to separate the "private" and the "service" spheres in the ARb software.

KEYWORDS

agent robot, "making sense", learning in a group, the ability to "read and write", "reflections", "private life"

1. INTRODUCTION

The goal of our work is to enhance the cognitive abilities of an agent-robot, a model of "reasonable man" that must be able to "understand the sense of texts" and to "think" in natural language. For the creation of such a model, it is rational to integrate two systems - two models corresponding to two modeling objects the theory and practice, of whose creation develops, to some extent, separately. These objects are human reason and the human body proper, and the models are, respectively, an agent and a humanoid robot. In the first case, we create a "thinking" system, that "has abilities which are traditionally attributed to human reason", while in the second case, we create a "mechanical" man who "understands and follows orders" [2]. The integrated model will already be a model of Homo Sapiens. Integration must give an additional positive effect to the cognitive development of an agent, in particular, to facilitate the reflection problem thanks to training in a collective of analogous agents and humans, as well as to using visible analogies with the actions of humans.

It is assumed that the functions of a human brain can be performed by a computer with a video sensor and an image recognition unit, as well as an acoustic sensor, in spite of the fact that there

is no comprehensive idea of the mechanics and operation of the human brain at present. In conformity with the goal stated above, we shall set the task of teaching the proposed agent-robot (ARb) from the start of its life, like teaching a child. ARb must learn to "read" and "write" and thus get an opportunity to continue its education at a later stage. It is proposed to refer teaching an ARb from "childhood", and educate it in the future to the ARb's "private life", independent of the work for which it has been constructed. Allocation of a part of an agent's functions, i.e. a part of its software, as "private life" was proposed in work [1] that considered the method of evolutionary development of agent in a "population" (MAS) based on the imitation of their "sexual" reproduction. Evolution was achieved thanks to gender division of agents that were identical in all other respects, and "pairing" them for creating "progeny". The probability of a positive "mutation" was provided upon the "birth" of any offspring, and there was a competitive struggle for the creation of offspring, i.e. "natural selection".

This article considers only teaching an agent-robot, a model of a human being from the moment of its "birth" until it reaches independence at the "level of a schoolchild".

An agent-robot must speak a natural language in order to "think". Thinking is impossible without language: a child who has not gained speech skills and has not learned to understand words in infancy does not master the basics of logical thinking and actually cannot become a normally thinking human [3,4]. Many works stress a close correlation between language and thinking [5,6]. The outstanding linguist L. Bloomfield [5] specifies this explicitly: "Talking to oneself or thinking". Surely, healthy humans do not "talk to themselves" (aloud) but they discourse "inwardly", using words, notions, as well as visible images - and, to a lesser extent, audible and other images. In our model of a human being, an automaton must master a natural language (NL) and, using mainly visible images, get a notion of real ambient objects and phenomena in the form of "written" and "read" words, phrases and notions.

1.1 Basic Provisions

The model of a "homo sapiens" must imitate the functions of a human organism, brain functions first and foremost, to perceive the environment, to adapt to it, and, last but not the least, to discharge the principal function of all flesh: reproduction [1]. Work ("service") requires an ARb to be able to solve complicated computational and logical problems, *including creative ones* [2,7]. Where upon the intellectual achievements of an ARb "in service" will influence the overall development of an ARb and its "private life" positively.

Recent monographs and course books (e.g., [2,7-10]) consider various methods of creating artifacts that have an "artificial reason", whereupon human reason, consciousness, and the brain are modeled separately from his or her bodies. This corresponds to the definition of AI as a science "studying human mental capabilities with the help of computers" [2] but does not aid the building of a generally adequate model of a homo sapiens.

The words "sense", "to understand", "to realize" in the summary and hereinafter in the article are put in quotes, since in respect of an Arb, one may actually say only that an ARb perceives texts, including illocutionary acts that correspond to reality. The reality of the objects and phenomena, actions and concepts observed by an ARb and described in the texts processed by it is determined because the text matches the images obtained by the ARb from the environment (the environment may be virtual), from video sensors, and the sounds communicated by audio sensors. All static and dynamic images of objects and phenomena and the sounds corresponding to them are recorded in the memory of an ARb. The match of a record with reality is also confirmed by information exchange and coordination of an ARb's actions in a collective of other ARbs and

humans. In order to verify the reality of words, including abstract ones, and then also the reality of notions and phrases, an ARb (like a human) reduces the whole text to the words and concepts whose reality has been learned by the ARb before, with the help of images and sounds. Before pronouncing a statement, a human often plans the content of that statement in the form of images connected with each other. For example, when a human answers the question: "How can I get to (somewhere)?", he or she usually describes the route he or she "follows in his or her mind". This example demonstrates the possibility of coincidence of a visual image (scene) with the semantics of a text in natural language.

The proposed method of teaching a language to an agent is based on a hypothetical concept of the mechanism of human meditations. It is assumed that the memories in a human brain accumulate throughout his or her life information not only directly in the form of texts, oral and written statements perceived by him, not only because of images (scenes) seen and perceived by him or her in the environment or in the form of static and dynamic images on various media, but also indirectly. That is, a human can also memorize notions obtained from texts he or she has heard or read or even fully imagined [11]. An independent speech or text unit - a phrase (sentence) can be compared with images in photographs and/or films containing the semantic meaning of that phrase, "a graphic phrase". Whereupon a phrase (one word or a number of word) and a respective number of photographs or a film are superventive, they determine each other. Great importance is attached to superventiveness of consciousness (texts) in human brain and written texts in some language in modern philosophy [12]. If we think, according to [12], that a text phrase in human consciousness is a mental phenomenon, then the respective visual images are perceived by the brain as physical phenomena. But it is possible to recognize visual images (including, e.g., images seen in a dream) as mental ones, and consider their subsequent presentation as a physical phenomena. It is the relationship of superventiveness (a phrase consisting of words - a graphical phrase) which is taken as a basis of mastering a natural language by automaton in our work.

The creation of a cognitively developed model of a human being is also implemented by teaching an ARb in a collective of ARbs and humans using methods similar to those of teaching a child. The "preschool" education course is primarily intended for mastering the communicative function of a language. Knowing a language (or languages) will allow an ARb to master the discursive method of "thinking" later [3, 4, 13].

2. BUILDING AN AGENT-ROBOT

A modern agent can be used as a basis for creating the software of an agent - the "brain" of an ARb, while existing humanoid robots may be used for designing and programming the "body" of an ARb (a robot) [2, 16, 17]. The "body" of an ARb - an anthropomorphous robot (an autonomous robots from the "mobile robots" class) reproduces approximately the functions of a human body [2]. In particular, an ARb can walk, carry objects, sit down, etc. [18]. A robot as a part of an ARb must reproduce a holistic organism similar to a human externally, and by some functions it carries out; it is necessary to increase the likeness between the modeling object and its model. The structure of an agent was proposed in [1]. Its main structural unit is the supervisor (similar to the assumed structure of a human brain [11]) whose main algorithm is the target function (hardwired) - the function of reproduction ("instinct").

It is assumed that an ARb will develop cognitively - "grow up", like a human. For that purpose, it is proposed to teach the ARb since the moment of its "birth" and until it receives (beyond the framework of this article) education, similar, e.g., to human university education.

The role of the existing AI systems - agents built as computer software - still remains the principal one in the technical implementation of the proposed model. The potential of the causal

dynamics of an agent, like other automatic creations based on modern computers, is high; in principle, it is the same as that of a human brain [11]. In particular, the speculative powers of a human being are directly connected with his or her memory; an ARb can be built only using all kinds of memory (random access memory, permanent memory, etc.) and modern advanced methods of loading information into them and its extraction. The data storage media of an ARb must contain texts, symbols and sounds, as well as static and dynamic images of the environment (image memory, and video random access memory (RAM)) [19]. We should note that "memory is the dominating function in human thinking process" [20].

An ARb must have all necessary sensors. The functions of the video sensor are performed by optical devices and specialized processors that are necessary for recognition of an image and its presentation [2, 21]. The audible information about objects obtained from audio sensors is an important additional characteristic, confirming the correspondance of an image to its verbal description, as well as determining certain covert features (e.g. a motor is running/is not running). The issues of speech recognition by robots and acoustic models [21] are not considered in this work.

An ARb communicates with a collective of fellow robots and humans only in the written form, "converses" and "thinks" using "oral written speech" [22]. An ARb perceives written texts, but also oral speech (apparently, with significant limitations (phonetic)).

3. TARGET FUNCTION OF AN AGENT-ROBOT

It is rational to set the life purpose of agents proposed in [1] - "reproduction" as the principal target function of an agent - robot (ARb). The programs imitating sexual instinct are inaccessible to an ARb for change during its whole lifetime. In [1], competitive (natural selection) is imitated in the process of reproduction, and every agent must achieve the best performance figures in work, as well as in the "bringing quality "posterity" in the world". Successes in the achievement of these goals were evaluated in bonds (an analogue of money). Such an evaluation is sufficiently universal; in particular, it can be interpreted as a function of utility. It is better to use numerical score for learning, as is customary. Extensional increase in the knowledge of an ARb can also be scored. Work (service) and therefore, the usefulness of an Arb, can also be scored by its employer (as well as its teacher); the score may be equivalent to bonds or even directly proportional to money. Thus, the implementation of an ARb's life purpose is expressed specifically - in getting the maximum score. When a certain score ("a secondary school diploma") is achieved, the principal target function starts acting. At first, a personal goal or task is formulated for an ARb by its teacher, and then the ARb is guided by its own experience: it chooses specific goals in view of the score it can get upon their implementation.

4. KNOWLEDGE BASE

A special Knowledge Base (KB) is created in the memory of an ARb. In addition to the generally adopted KB [23] of an agent, it includes a set of special rules and algorithms, as well as a guide to natural language (NL). The guide contains a "Great Definition Dictionary" including all the words of a certain NL, both significant and functional ones. Every word in the Dictionary and all its grammatical forms are explained by other words, like in a generally accepted Definition Dictionary [24], whereupon the most frequently encountered adjectives are mostly given for subject nouns, and necessary combinations of significant words with function ones, for various grammatical forms.

The significant words that may be provided with video images are presented with the respective images, static (mainly nouns (Noun Phrase - (NP)) and/or dynamic - films (mainly verbs (Verb Phrase (VP)) It is rational to use images of objects and phenomena from the Dictionary during image recognition using the extensional method, the method of comparing with the prototype [21], whereupon, as in a human being the modeling object, in our model the optics of the video sensor is directly linked to the agent, the "brain" of an ARb. Searching for a prototype is facilitated thanks to the fact that the location of an ARb is known.

The Great Definition Dictionary has sections: the *alphabetical* one, where words are arranged strictly in alphabetical order and the *classification* one, where words are distributed in a certain order, depending on their meaning. In the Alphabetical Section, all words are provided with definitions, images and films for every subject word, sounds, as well as address indexation: a reference to the coordinates of every word in all other subsections of the Dictionary, whereupon indexation must be reversible, i.e. link all the sections (subsections) of the Dictionary together.

The background knowledge is presented in the principal part of the KB. The Dictionary contains only explanations intended for "primary" language teaching of an ARb, pointing out the attributes and links being the key difference: the cause of referring the word to this class, subclass, etc. Sufficiently complete characteristics must be given for the words and concepts, meaning the objects and phenomena of the animated and the inanimate nature. For artifacts: scope, purpose, and principles.

Homonyms are defined by separate explanations for each meaning, and the respective images and/or sounds. Abstract words (e.g., abstract nouns: work, play, friendship) especially need special detailed explanations. Images and films are also used in explanations of abstract words. For example, a film: humans and robots carry tables and chairs from one room to another (caption: "Work"), people build a wall and lay bricks ("Work"), etc. Antonyms may be used as well, e.g., the image "work" - "rest" or by sound: birds "sing" - "keep silent", etc. When explaining abstract words which are complex for an ARb student by object words, the principle of reduction is used. Correction in case of an insufficiently exact reproduction of an abstract word is performed after a more complete study of a language by the ARb [4,22,25]. A number of abstract words whose semantics are connected, first and foremost, with the human community (e.g., vexation, pity) can be mastered by an ARb only after it gets sufficient education and "life" experience.

The classification section has a hierarchic structure in the form of a word search tree [2] using images. The top ontologies are "Nature" and "Artifacts", the next level of the "Artifacts" section is "Cities", "Villages", etc. All images (photographs and films) correspond to the names of objects and actions and are explained by their respective characteristics, as well as individual images of accessory objects and phenomena that are also named. The words have attributions and are linked by indexation with other sections of the Dictionary.

All static and dynamic images (all states) down to the last leaf of the search tree must strictly correspond to the names to which they pertain.

A complete grammar of the NL is presented in the Dictionary: spelling and punctuation with examples in which words also must be indexed. There are also links to programs of moving from one explanation of a word (first and foremost, an abstract one) to another explanation of the same word; the chain of words is continued until all the cited words are explained by images and/or sounds - e.g., one of the meanings of the abstract word "choice" [24]: "Choice of one from two or several objects..." then "Choose - Select", "Select - take, single out from the total number" and, finally, "take - capture by hand" with a respective image. Rules (e.g., the rules of conduct, the safety rules) and other words, e.g., commands requiring necessary actions from the ARb must

have references to the algorithms intended for these actions - e.g., words - commands (in the imperative mood): "go" or "read" are received and executed by the "brain" of the ARb. In case of words - ARb commands executed by the robot proper, e.g., "raise an arm" or "sit down", robot algorithms (not considered by us) must operate.

The reality of a text is confirmed by comparing a phrase of the text with a pictorial phrase that must have a strict concatenation of photographs and films, whose semantic meaning coincides with the respective words from a phrase in a text, whereupon for concatenation of a pictorial phrase, active objects presented in photographs must be singled out in films.

5. KNOWLEDGE BASE OWN

From the very beginning of its life, an ARb must gather its own experience of reducing words and phrases to reality. The initial perceptive abilities of an ARb appear at the moment of its "birth" and then the whole totality of observed objects and phenomena, a universe of scenes, is remembered.

All the used and learned words, phrases, verified by conformity to reality, pertaining to objects and phenomena both confirmed by images (static and dynamic) and sounds, and those with no video and audio images are recorded in a separate memory – their own memory that also contains images and sounds. The contents of this memory are individual for every ARb. A Knowledge Base Own (KBO) must have its own Definition Dictionary and the same sections whose structuring and indexing is similar to the one in the principal KB, including the Classification Section of images with an hierarchic structure - the word search tree. A KBO may have a bidirectional search tree [2] where after discriminant analysis, lower ontologies are created on the basis of clusterization methods- e.g., words with new attributes are accumulated gradually in the section "Tables", subclass "Furniture": a table on four legs, on three legs, on one leg (pedestal), a writing table, a dining table... and a special agent algorithm creates a cluster, e.g., "Tables on 4 legs" and adds it to the general classification structure.

A KBO has additional sections: Common, Personal and Working. The first two of them contain the same sections received by a ARb from its parents' KBOs [1], i.e., parents' joint KBOs form the basis of their offspring's KBO. Inbreeding accelerates and improves the creation of an effective offspring ARb's KBO. Gradually, a subjective "internal model of the external world", the "embryo of consciousness" is formed in the KBO. The KBO of an "adult" ARb with addition of the algorithms that were created or "modernized" by it form a *personality*. This is an artificial personality (AP), one and only, like a human personality. Unlike a human, the whole anthropomorphous part of an IP is a robot; besides, some units of the computer forming the ARb may be removed or replaced, whereupon the IP remains operational (the ARb is "alive").

6. TEACHING AN AGENT-ROBOT

Modern methods of cognitive development of a child [3,13] may be used as an analogy for the basic teaching of an ARb. The principal thing is learning a NL: words and concepts. Everything an ARb sees and "learns", all the objects and actions in the environment become the content of its KBO. We should note that the environment for an ARb at the beginning of its life is a room (in a laboratory), then a building, a street, etc.

Images received by the ARb from the observed environment (including the virtual one) are dependent in each specific case on the shooting angle, the lighting, etc. Therefore, on different occasions, images of objects and actions may differ from one another even in the memory of one

and the same ARb. Even more so, different ARbs must have such differences, i.e., the "personality" of every ARb is individual and unique. In the process of "reading", images of objects and phenomena from the sections of the Dictionary are compared with the images actually perceived by the video and audio sensors. We should note that all the sets of objects perceived by an ARb are finite, but some of them can be very large (e.g., the list of images and names of animals) and therefore, they must be replenished throughout the lifetime of an ARb. An important stage of learning is the appearance of reflection in an ARb: the ability to pay attention to the activity of its own self in relation to ambient objects that is attained by using special methods in a child as well [4,13,20]. The classification of words with their attributes by categories and sections both for the world of nature and for artifacts is available in the KB; it is transferred to the KBO along with mastering words, phrases and concepts, creating the "language of an ARb". The agent "brain of an ARb" belongs to the type of "knowledge-based agents" [2]. Its KB contains background knowledge, while a supplementary function of the KBO is to obtain knowledge from experience in the process of solving current problems as well.

We shall present informal algorithms of the functioning of an ARb in pseudo-code, interpreting it in accordance with [2,7]. PROLOG, LISP programming languages, and their modernizations may be used for the implementation of ARb algorithms written in pseudo-code.

Let us cite a common scheme of an algorithm (1) of a KBO query (data input and request) [2]. The ARb transfers the perceived results (percept) to the KBO (knowledge base own) and returns the action developed by them (action). The ARb also supports the KB and the KBO by its "conscious experience" [12]. The perception of the ARb is entered in the KBO with the continuous help of the Tell operation. Additionally, the data of its own problem solving experience (data aggregate) are entered.

A request to the KBO about actions is provided with the help of an Ask operation. Answers are registered thanks to the Tell operation, and an action is implemented. Problems and ways of solving them remain in the memory, and the ARb will take them into account in the future as its own experience.

```
funktion KBO-ARb (percept, state, data aggregate, problem)
returns action
inputs: percept, problem, data
static: KBO, state
Tell (KBO, Make-Problem-Words (percept, data aggregate, problem))
problem := Ask ( KBO, Make-Action-Query)
Tell := Tell (KBO, Make-Action-Words (action, complement))
return action
end
```

Here, the Make-Action-Query function forms a statement: a request, whose action must be performed by the ARb in this case, and the Make-Action-Words forms a statement confirming that the chosen action was performed. The Tell and Ask functions contain the necessary mechanism of logical inference [2].

Then (by default) it is assumed that the above scheme of knowledge based query is used in all algorithms upon interaction of the ARb with the KBO and the KB and transmission of results from the KBO to the KB for the verification of conformity of images to words.

An important factor for studying a language is communication (in this case, written) in a collective of individuals who know it, the "language environment" [6,27,28]. An ARb must

master a language by “collecting” during studying in its "adult" life as well, e.g., as an element of MAS-ARb [15].

The task of a teacher is to help the ARb to memorize and "master" the first 200-300 words and the grammar of a NL, to enter *at a first approximation* a visual and lexical representation of "concentric" world structure into the ARb's memory: the sky (the universe in the most general terms) and in more detail: a city, a street, a house, a room (laboratory), as well as an ARb, a human, a collective of humans, etc. Another task of a teacher is to establish a dialog: Teacher to ARb, ARb to ARb. The reflection of an ARb and link between words, concepts, ideas in the KBO with a section of the Dictionary is needed by the ARb in searching for new words and algorithms "linked" to them, including algorithms of the actions undertaken by the ARb. An important task of a teacher during teaching abstract words to an ARb is the ARb's mastering the use of the following modal verbs in texts: to wish (want) and to be able. The verb *to wish* corresponds with the "need to have" [3,13]. Let us cite an example. The teacher's task: to organize a didactic game with the ARb mastering reflection and the verb "to wish" in the game. Four ARbs play. The teacher gives them each a pack of photographs of animals. The ARbs open their photos in turn. The one that finds the desired thing, e.g., a giraffe, raises its hand and reports (writes on its display, located, e.g., on its "forehead"): "Me!". After having been repeated many times, the game must result in the automaton's mastering the concept of reflection. Once, as the teacher gives the decks of photos to the players, he or she "forgets" to give a deck of photos to one of the ARbs. Observation may bring the "deprived" ARb to apply to the teacher. The ARb sees that other players have packs of photos it also needs to have. It "wishes" the pack. The "application: of an ARb, e.g., not only a statement with a request but also a gesture (an outstretched hand) is an important positive result. All the ARbs see the gesture and perceive it as the image of the word "Ask", "Wish", "Want" [13] and the teacher may install the algorithm "Request" (or activate the one available in the KBO).

The ARb (in a similar way to a child [13,27,28]) must imitate (without analyzing the phenomenon it observes) humans and robots and, first and foremost, the teacher- e.g., a human or an ARb of the same age as the taught ARb raises his/its hands up, and the ARb also raises its hands up. The teacher walks and the ARb walks.

An example of the imitation algorithm (2).

The goal of ARb1 is imitation of the actions performed by another ARb. ARb1 observes the actions of ARb2 in time, translates images into words, finds algorithms corresponding to this words in the KBO and repeats (repetition) the "task" (job), i.e., all the characteristic movements (movements) of ARb2 in time (t).

funktion Imitation (repetition movements) **return** aktion (t)

inputs: perzept

static: goal repetition movement

state environment, home position

t - counter time *from zero* ($t_0 \dots t_i \dots t_n$)

goal := repetition (job,t)

aktion := First repetition (t = to)

return aktion

action := Tell (KBO, movement) *KBO query*

t := $t_0 + 1$ repetition

return aktion

action := Tell (KBO, movement)

action := repetition (t = $t_i + 1$)

.....
return aktion (t)
end

From now on, if the action is performed by the robot proper (the "body" of an ARb), our algorithms contain only a command to the robot's control unit, the programs of control of robot components (manipulators, pediculators, etc.) are not considered.

The ARb supplements perceive words: nouns (Noun Phrase - (NP)) and verbs (Verb Phrase (VP)) by extensional attributes, characteristics of objects and phenomena: ANNP AVVP where AN and AV are attributes. With regard to these attributes and comparing the set of available examples, the ARb uses an algorithm that replenishes the classification sections of the KBO, distributing all words by the respective classes, subclasses, sections, etc. The list of word-to-image conformities in the KBO is continuously replenished.

The verification of the words received from a video sensor and meaning of an object or a phenomenon (Search word image matching (SWIM)) [2] is performed by the ARb by querying the images in the KB Dictionary. When searching for matches, the upper ontologies of the search tree and the closest subsequent sections in the beginning of teaching the ARb are missed because the perception of the ARb gives an image of circumstances at the ARb's location; this corresponds to the bottom search tree levels. An element in the state space is compared with a similar one found in the ordered set of images of the bottom tree level, whose elements (members) form a list of images. The range of search for matching images is also determined on the basis of the Classification Section of the Dictionary. The SWIM algorithm uses a recursive procedure. At first, this procedure compares the observed image with the first image in the list of images in the Dictionary at the chosen level. In case of the failure, the ARb goes over to the next image, and the first member of the list is removed. If a result has been obtained, i.e., it was found that the images are matching, this is successful completion of the procedure and the whole algorithm. After successive comparing of the whole list without a result (all elements have been excluded from the list), the search is transferred to another classification tree subsection that is the "closest" by way of its attributes.

Algorithm of verification of the words obtained from the video sensor (3)

function SWIM (name image) **return** success (Tell KB word) or failure

inputs: percept (environment, object (image))

static: goal (fix matching image name word)

state (environment, objects, classification tree KB)

if member (first matching image) result **then return** success

end

else member (first matching image) empty

then return failure

.....
tail := member (matching image) tail is *the list after removal of the first member (tail of the list)*

.....
result **then return** success

end

end

In the process of learning, the ARb increases its stock of words and phrases in the KBO, and must reach a level that allows it, after syntactic analysis [25-27] of a received image, to establish the order of words in the sentence, comparing these words with images in the same order to the master, i.e. to confirm the reality of the semantics of a message. Whereupon, searching for the

meaning of abstract words requires a number of transitions from word to word, and then to the image that confirms (or refutes) the reality of a message.

Comparisons and detection (after sequential consideration of a set of examples and their analogies with those already mastered and checked by "own practice" images (films), whose description is already available in the KBO) allow the ARb to find a match between the target function, and the general motivation of phrases: the principal factor for semantic identification of a newly received message. This looks like a version of the inductive method of teaching [2, 7].

7. COMMUNICATIVENESS

Written messages in a natural language or languages are used for the communication of an ARb with its other ARbs and humans, therefore, a learner ARb must be able to "read and write". The step-by-step algorithm of "reading" a message by an ARb begins with a syntax analysis of the text of the message [25, 27]. After that, the ARb carries out a search of images in the KBO for ANNP & AVVP words with the same or similar attributes. Then the ARb replaces attributes in images of the NP and VP found in memory by those present in the text if it is necessary and possible. After that, it builds consistently a complete image (pictorial phrase) according to the text of the message. The last step (going over from images to the words of the whole text) is feedback.

Let us write the solution of the "reading" task in pseudo-code (4).

```
function Read ARb (text) returns result
inputs: text message
static: goal (set words text – mapping image)
words text
perzept image
text:= text parsing perform: text parsing
function Search Mapping (words image) returns success search for matches
search := Make-search-image (words-queue NP VP) search word = image
substitution := Exchange attributes (ANNP AVVP, KBO) exchange attributes
make integration := Images (full text) ("understanding" the written in the seen)
transfer := Images-words (full text) control: feedback
Images (full text):= Read ARb (text)
return result
end
```

The feedback in the last step of the algorithm is the operation of writing a text according to the "visible scenario" created by the ARb: a film reflecting the message text.

Let us write the solution of the text "writing" task in pseudocode (5).

```
function Write (text) returns result
inputs: perzept image
static: goal (mapping image - set words ) goal: images match words
perzept image
search := Make -search-words (words, text, images)
make integration := Text (full images) writing the full text
("the seen") in words
make-write := Text (full text)
returns result
```

Thus, an ARb is already able to communicate: to read and write.

In the general case, every instruction (command, directive), e.g., of the teacher, requests, e.g., of another ARb and its own planned actions for a learner ARb, can be written down in the form of a verb in the imperative mood, e.g., those necessary for communication: read!, write!, go!, look!, whereupon the sign ! (or another one) confirms the command. A verb in the imperative mood is a command.

The algorithm of execution of commands, own decisions or requests by an ARb (6).

```
function Command (execution) return action
inputs: command
static: goal (execution command) .
state (environment, objects)
search := Make-search-image (matching words name-image )
matching := word - name image
action := execution command
return action
end
```

During "initial" instruction, an ARb executes only elementary algorithms. Its autonomous actions may be based on the algorithm of imitation, ready-made game algorithms, direct instructions of the teacher, etc. For example, in imitating a human, according to algorithm (2) an ARb wiped its feet upon entry to a room. Next time, it can execute this procedure already using its own record of the sequence of actions in the KBO as an algorithm as the pattern for executing these actions. In another case, an ARb changed its footwear when it entered a building and left the outdoor footwear in an allocated place. As a result, it can use either this or that algorithm depending on the circumstances, without synthesizing a new algorithm like: **If** enter into a building from outdoors, **then** must wipe my feet **or** change my footwear. An ARb is taught logic (first the "everyday logic" and then scientific logic) at a later stage.

The achieved abilities: *to read - see the image of objects and phenomena matching the semantic meaning of a message and to write - describe the seen things with words allow a learner ARb to communicate with other ARBs and with humans.*

An example of communications and actions. An ARb gets a message: "Lotto game in room 512. Come!" The ARb uses the "reading" algorithms (4). The message text is divided into three images, three scenes: "Lotto game/ in room 512./ Come!" After "reading" the message and the command (invitation), action planning is necessary[2]. The ARb begins planning from the search of images and words, search of algorithms, search for precedents: its "conscious" experience in the KBO. The ARb's plan of going over from the ground floor to the fourth one contains five images (films): Go to the right in the corridor/ enter elevator No. 3/ leave elevator on the fourth floor/ go left/ to room No. 12. To execute the command: "come!", the ARb uses a (6) type algorithm. (Robot's actions: makes a step first by its left leg, then by the right one; presses a handle and opens a door... etc. - should be classified as algorithmic robot software [16,17]).

The informal algorithms provided by us are only examples. Many more algorithms will be required even for the "preschool" education of an ARb: refining a specific goal, priority of actions executed by an ARb, etc.

Hundreds of algorithms out of the thousands that will be required to continue the education and other actions of an ARb must be installed during the "elementary" education.

8. ARTIFICIAL PERSONALITY (AP)

Further education of an ARb and its communication in a collective must broaden its perspectives, erudition, and the circle of the ARb's "friends" among its fellow learners and humans [28,29]. Like a human, an ARb can be considered as a *personality* when it has acquired social qualities in the process of socialization, for whose goal the ARb must be in a collective of other ARbs and humans. Time must be allocated to an ARb periodically in order to make it possible to establish friendly (personal) ties. Every AP will be characterized by a certain activity, the grade of independence, its own attitude to the external world (the environment) and the attitude to itself. In other words, the artificial personality AP (of an "adult" ARb) will already have a self-consciousness, a "psyche". The above characteristics form the basis for the specialization of an AP, its education, and the circle of its business and private contacts. If an AP works in a collective consisting mostly of humans, e.g., in a bureau or an institute, the combination of an agent with a robot is convenient and aesthetic. If an AP works with machines, e.g., as an electric car driver, the combination of an agent with a robot is irrational and the robot can be retired.

An anthropomorphous robot is necessary for the formation of an artifact as an artificial personality. However, afterwards it can be considered as an important, but only one of many possible actuating mechanisms of the automaton that is used, e.g., for the direct replacement of a human at his or her workplace.

During "elementary" education, an ARb may master two or three languages at the same time. The names of one and the same object in different languages will allow an ARb to determine the names of observed objects and phenomena or to describe a scene more reliably and more precisely. For example, mastering two languages and the availability of detailed NL 1 - NL 2 and NL 2 - NL 1 dictionaries will facilitate in an ARb the search for a match between words and concepts and the most difficult search for matches for abstract words, which will ensure the mastering of the meanings of homonyms, and influence the erudition and area of thought of the ARb positively. Mastering languages may be necessary to the AP of an ARb on the "job" for "machine" translation as well, and if the ARb is highly intellectually developed, e.g., for taking part in international scientific conferences.

9. CONCLUSION

This article proposes a new approach to enhancing the cognitive abilities of an intellectual agent robot (ARb), a model of a *homo sapiens*.

A new approach presumes an analogy in the development of an ARb and a human, from the moment of the "birth" of an Arb, until it evolves into a "personality". It is proposed to construct a model of a *homo sapiens* as an intellectual agent, integrated with a humanoid robot. This integration is necessary from the beginning of the agent's "life", as it is educated in a group of ARb's and humans, and is afforded an opportunity to imitate humans directly. The socialization of an ARb begins in a group of students.

The meted of constructing the agent proper, the brain of an ARb is consisted of replacements of individual organs of a human brain/head, categorized by their functional purpose, by each respective artificial organism with analogous functions. In particular, obtaining and memorizing images is a function of the brain. Like a human, an ARb memorizes images obtained using video sensor optics and an image recognition system, and compares it to images saved in its knowledge base (KB). For mastering a natural language, the standard KB of an agent is supplemented for an ARb by a "Definition Dictionary of a NL" with images of objects and phenomena in the form of photographs and films. The dictionary consists of several sections, including a Classification

Section in the form of a search tree representing the living environment of the ARb. Classification facilitates the verification of matching of words to actual images obtained from the video sensor.

The matches mastered by an ARb: a system of images (photographs and films) and a system of words (a text) make it possible to consider a text phrase in a certain NL, and a set of images (static and dynamic), i.e., a "pictorial phrase" as superventive systems, confirming the reality of a text. All matches are entered by an ARb into a Knowledge Base Own (KBO), that must have a dictionary with the same sections and the same structure (tree) as the dictionary in the main KB of the ARb. The KBO of an ARb includes images, sounds and words corresponding to them, as well as algorithms "inherited" from its "parents" [1]. A model of the outside world specific for the ARb in question is created in the KBO.

It is proposed to use a NL for the communication of ARb's between themselves and with humans, and for the "reflections" of an "adult" ARb. In compliance with the adopted analogy with a human life, immediately after the "birth" of an Arb, it starts to learn a language under the tutelage of a teacher, whereupon the ARb communicates with humans and other ARb's. Knowing a NL predetermines the ability to read and write for an ARb. Written communication replaces oral communication for an ARb in a group of fellow ARb's and humans. An agent robot communicates and later "reflects" using a NL. After learning a language and completing general "primary" education, an ARb undergoes special education needed to fulfil "service" tasks.

The function of "reproduction" involving the production of numerous high-quality progeny, as proposed by us earlier, [1] is adopted as the "life purpose" of an ARb. The implementation of the "purpose" consists of rating an ARb throughout its life using points during education, and bonds when "in service". The total value of the bonds (and points) characterizes the "track record" of every ARb, and predetermines its choice of the best partner for creating progeny, i.e. ensures "natural selection" and evolution in the community ("population") of ARb's. For implementing its "life purpose", an ARb that has completed its education is to have a "private life" in parallel with its work ("service"), just like a human. A separate "private life" of an "artificial human" allows agent robots to improve their own cognitive development and general erudition independently of the work they perform, and to refine their own world outlook. A personal vision of environmental objects and phenomena and the "genes" inherited by an ARb from its parents shape the individual artificial personality (AP) of an ARb.

The new approach to constructing an ARb, the model of an "artificial human", brings substantial positive results. An ARb: 1) is proficient in a NL, like a human, on the basis of remembering the names of the real world objects and phenomena it has recognized; 2) communicates in a NL in writing with other ARb's and humans, starting from a group of ARb's and humans during primary education; 3) has a self-consciousness – individuates its own self in a group, thanks to special teaching methods; 4) "understands" the sense of texts and words in a written dialog on the basis of presentation of worded objects and phenomena designated by words; 5) can read and write. Elementary algorithms of these (and other) actions are based on understanding (according to 4)) the sense of words; 6) perceives experience, knowledge and relevant algorithms "inherited" from its parents [1]; 7) has a "life purpose" and therefore, strives for education and pursues a career [1]; 8) has a "private life" since it is afforded a certain amount of free time daily in parallel to work, and improves the quality of its work throughout its "life"; improves its own cognitive abilities substantially.

REFERENCES

- [1] Stalsky V. V., 2014 Evolyucionnyi metod kognitivnogo razvitiya obuchayushchihsvya avtonomnyh agentov. (Evolutionary Method of Cognitive Development of Learning Autonomous Agents). //

- Iskusstvennyj intellekt i prinyatie reshenij (Artificial Intellect and Decision Making), No. 2, pp. 52-62 (in Russian).
- [2] Stuart J. Russel and Peter Norvig, 2003, Artificial Intelligence. A Modern Approach. Second edition. Prentice Hall, NY: Upper Saddle River.
- [3] Smirnova E.O., 2006, Detskaya psihologiya (Child Psychology). Moscow, Vldos Publishers, (in Russian).
- [4] Ermakova E. S., 2007, Razvitie gibkosti myshleniya u detej. (Developing Flexibility of Thinking in Children). - Saint Petersburg, Rech Publ.(in Russian).
- [5] Bloomfield L. 1933. Language. London: George Allen and Unwin.
- [6] Maslova V. A., 2011, Vvedenie v kognitivnyu lingvistiku (Introduction to Cognitive Linguistics). Moscow, Nauka Publ. (in Russian).
- [7] Luger J. F., 2005, Artificial Intellect: Strategies and Methods of Solving Complex Problems, Williams Publ.
- [8] Gorz G., Schneeberger J., Schmid U. (Hrsg.), 2014 Handbuch der kunstliche Intelligenz. (5.Aufgabe), Munchen: Oldenburg Verlag,
- [9] Yasnitskiy A. N., 2010, Vvedenie v iskusstvennyi intellekt. (Introduction to Artificial Intellect. Moscow, Akademiya Publ., (in Russian).
- [10] Romanov M. P., 2008, Iskusstvennyi intellekt i intellektualnye sistemy. (Artificial Intellect and Intellectual Systems). Moscow, Nauka Publ., (in Russian).
- [11] Rapoport G. N., Gerts A. G., P.1 - 2005, P.2 - 2011, Biologicheskij i iskusstvennyi razum. (Biological and Artificial Intellect]. Moscow, URSS Publ., (in Russian).
- [12] Chalmers David J. 1996, The Conscious Mind: in Search of a Fundamental Theory. Oxford University Press.
- [13] Volkov B. S., 2000 Doshkolnaya psihologiya (Preschool Psychology). Moscow, Academ. Project Publ., 2000 (in Russian).
- [14] Grosan C., Abracham A., 2011, Intelligent Systems. A Modern Approach. Berlin: Springer.
- [15] Büther R., 2010, Automatisierte Verhandlungen in Multi-Agenten-Systemen. Wiesbaden: Gabler Verlag.
- [16] Hertzberg J., Lingemann H., Nüchter A., 2012, Mobile Roboter. Heidelberg: Springer.
- [17] Bekey G. A., 2008, Robot Learning Autonomous Robots. Massachusetts, London: The MIT Press Cambridg.
- [18] GuErin J., Thiery S., Nyiri E., Gibaru O. (2018) Unsupervised Robotic Sorting : Towards . . . Autonomous Decision Making Robots // IJAIA, Vol. 9, .2
- [19] Borkovskiy A. B., Zaichik B.I., Borovikova L.I., 1999, Slovar' po programmirovaniyu (Dictionary on Programming). Moscow: Russkiy Yazyk Publ. (in Russian).
- [20] Vygotskiy A. G., 1982, Myshlenie i rech. (Thinking and Speech). Collected works, V.2. Moscow, .Pedagogika Publ. (in Russian).
- [21] orsythe D. A., Pons J., 2002, Computer Vision: a Modern Approach . Boston: Pearson.

- [22] Kreidlin G. E., Krongauz M. A., 2004, Semiotika i Azbuka obshcheniya (Semiotics and the ABCs of Communication). Moscow, Flinr Publ., Nauka Publ. (in Russian).
- [23] Gavrilova T. A., Khoroshevskiy V. F., 2000, Baza znaniy intellektualnykh sistem (Knowledge Base of Intellectual Systems). Saint Petersburg, Piter Publ. (in Russian).
- [24] Ozhegov S., 2008, Tolkovyi slovar russkogo yazyka. (Explanatory Dictionary of Russian Language) .Moscow, Onyx Publ. (in Russian).
- [25] Marchuk Yy. N., 2007, Kompyuternaya lingvistika. (Computer Linguistics). Moscow, AST: Vostok Zapad Publ.(in Russian).
- [26] Sosnovsky B. A., 2005, Vozrastnaya psikhologiya. (Developmental Psychology). Moscow, Yuraye Publ.(in Russian).
- [27] Golitsyna O. A., 2016, Informatsionnye tekhnologii (Information Technologies). Moscow, Forum Publ. . (in Russian).
- [28] Skobtseva T. G., 2011, Kognitivnaya lingvistika. Kurs lektsiy. (Cognitive Linguistics. A Course of Lectures). Saint Petersburg University Publ.(in Russian).
- [29] Psikhologiya detstva. Pod redakts. A.Reana, 2003, (Childhood Psychology. Edited by A. Rean) Saint . Petersburg, Evro-znak Publ. (in Russian).