Algorithms of Natural Language Dialogue with Intelligent Robot NAO EVOLUTION

Anatoliy Shvetcov
Department of Automation and Computer Engineering
Vologda State University
Vologda, Russia
smithv@mail.ru

Vyacheslav Gorbunov
Department of Automation and Computer Engineering
Vologda State University
Vologda, Russia
gorbunov1945@inbox.ru

Sergey Dianov
Department of Automation and Computer Engineering)
Vologda State University
Vologda, Russia
dianovsv@mail.ru

Alexey Sukonschikov
Department of Automation and Computer Engineering
Vologda State University
Vologda, Russia
avt@vogu35.ru

George Rapakov
Department of Automation and Computer Engineering
Vologda State University
Vologda, Russia
avt@vogu35.ru

Kyril Kinyakin
Department of Automation and Computer Engineering
Vologda State University
Vologda, Russia
avt@vogu35.ru

2022 VI International Conference on Information Technologies in Engineering Education (Inforino) | 978-1-6654-0577-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/Inforino53888.2022.9782900

***Abstract*—This article is devoted to the usage of anthropomorphic robot NAO Evolution in teaching process of bachelors and masters at Automation and Computer Engineering department of Vologda State University. The paper discusses the content and implementation of software system of the dialogue man-robot on natural Russian language using the software package Choreography and dialogue creation tool QIChat. Fragments of software modules that implement the developed features are given.**

***Keywords*—anthropomorphic robot, natural language dialogue, human-robot interaction scenarios, semantic network, game module**

I. Introduction

Artificial Intelligence (AI) technologies are beginning to be actively implemented in industrial practice, with high government financial and organisational support (Russian Federation government decree no. 767 dated May 21, 2021). The University of the National Technological Initiative 2035 offers an extensive programme of activities for schoolchildren and students to form projects in the field of artificial intelligence, to provide citizens with additional professional education in the field of artificial intelligence and in related areas using the mechanism of personal digital certificates, to develop and update a digital solution for accounting and development of participants of communities in the field of artificial intelligence [1].

At Vologda State University, Department of Automation and Computer Engineering in the curricula of undergraduate and graduate training in 09.03.02 and 09.04.02 "Information Systems and Technologies", the scientific and technical direction of AI is represented by a number of interrelated disciplines, covering the main areas of AI and forming basic skills and abilities in the development and operation of intelligent systems [2].

In such disciplines as "Intelligent Systems and Technologies", "Distributed Intelligent Systems", "Infocommunication Systems and Networks" to implement theoretical knowledge and form practical skills and abilities to control and interact with technical intelligent systems the anthropomorphic robot NAO Evolution V5 of French firm Aldebaran Robotics is used [3].

Currently, anthropomorphic robots of this class are widely used in various universities around the world for educational and research purposes. In our country, NAO robots are implemented to solve scientific and educational tasks in such leading universities as the Baltic Federal University named after E. Kant. In our country, NAO robots are implemented to solve scientific and educational tasks at such leading universities as E. Kant Baltic Federal University, Kazan Federal University, N.E. Bauman Moscow State Technical University, St. Petersburg University of Telecommunications named after M. A. Bonch-Bruk.

The results of the research conducted under the auspices of the Kant Baltic Federal University, Kazan Federal University, N.E. Bauman Moscow State Technical University, Bonch-Bruevich Saint-Petersburg University of Telecommunications, Skolkovo Institute of Science and Technology, Tomsk State University of Control Systems and Radioelectronics [4].

According to a study performed under the auspices of UNESCO [5] a modern intelligent robot should provide such characteristics as mobility, interactivity (provided by sensors and mechanisms that collect relevant information from the environment and allow the robot to influence this environment), information exchange (provided by computer interfaces or voice and speech recognition systems), autonomy (i.e. the ability to reflect and draw conclusions independently and make its own decisions to influence).

The NAO autonomous programmable humanoid robot fully possesses these capabilities, allowing it to be used to solve and simulate a wide class of tasks traditionally within the field of AI.NAO H25 EvolutionV5 has the following specifications: height 58 cm, weight 4.3 kg, battery capacity 48. 6Wh battery provides 60 minutes of active battery life (90 minutes in normal mode), 25 degrees of freedom, 2 HD 1280x960 video cameras, 4 microphones, ultrasonic sonar, 2 infrared ports, 2 gyroscopes and 1 accelerometer, 8 foot pressure sensors, 9 tactile sensors (head and hands), communication via Ethernet, Wi-Fi. Human verbal communication devices provide speech recognition and reproduction in 19 languages, including Russian. The robot has two quite powerful processors, one in the robot's head - Intel Atom 1.6 GHz with NAOqi OS on Linux kernel, the other ARMv9 in NAO robot's chest (embedded Linux OS, compatible Windows, Mac OS, Linux, programming languages C++, Python, Java, MATLAB, Urbi, C, .Net,).

978-1-6654-0577-5/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: Universidad de Talca. Downloaded on June 27,2025 at 16:04:03 UTC from IEEE Xplore. Restrictions apply.

Using embedded software, the NAO robot can identify specific people, respond to voice commands, gestures and facial expressions of others, use expressive gestures to communicate, interact with users using speakers, microphones, cameras, tactile sensors, LEDs, hands, body movements. Built-in software modules include speech recognition and synthesis with voice control, it can recognise and remember faces and objects, and identify the nature and direction of sound sources. The extensive functionality of the anthropomorphic robot NAO allows considering it as a hardware-software base for research and development of models of interactive learning and self-learning in the 'Student-Robot-Tutor' (SRT) system. A necessary component for the implementation of such models is a dialogue in natural language between the robot as a subject of the educational process, and its other participants - teachers and students. The models and algorithms for organizing such a dialogue, presented in this paper, are a continuation of the research, the results of which were outlined in [6]-[7].

II. Structure Of Natural Language Communication System With NAO Robot

The project of communication with a NAO Evolution V5 Academic Edition robot in natural Russian language is a set of programs sent to the robot's memory, consisting of several basic modules connected by a central unit into a single whole. Initially, the Russian language is selected for the entire project so that it is possible to communicate with the robot in its native language. The user is then given an initial greeting and a choice of communication topics via the linking unit. The main topics for communication are the characteristic subject areas of verbal communication: weather, sports and robots. A separate block accommodates a synonym dictionary for the whole project, allowing for a variety of possible human-robot dialogues. Two additional blocks, game and function, form an extended verbal interaction with the robot, providing access to the control of physical functions that allow the user not only to carry out natural language dialogue with the NAO, but also to realize visual and tactile interaction.

The development and debugging of the natural language communication system was performed using the Monitor and Choregraphe utilities and development tools provided by the development company [8]. The built-in tool for development of complex dialogue systems QIChat, which showed sufficiently high efficiency, was used when creating the system. In the course of the project, the software code of standard blocks was complemented and modified, and a number of new blocks were created using the Python programming language. Access to the robot's memory for extracting and adding various data and files is made with Total Commander, which provides FTP connection. The main block diagram of the developed system is shown in Fig. 1.

In the process of laboratory works with the anthropomorphic NAO robot, students are mostly interested in debugging and testing verbal game interaction, where a dialog box containing an interactive game between a human and a robot, implemented in the form of a text quest, is investigated and tested. The game block also activates two games, one of which is associated with the recognition of objects whose images are formed in the knowledge base of the robot in the course of previous training on the presented samples, and the second involves movement to the identified objects by detecting and identifying objects by special tags, each of which transmits its digital ID number.

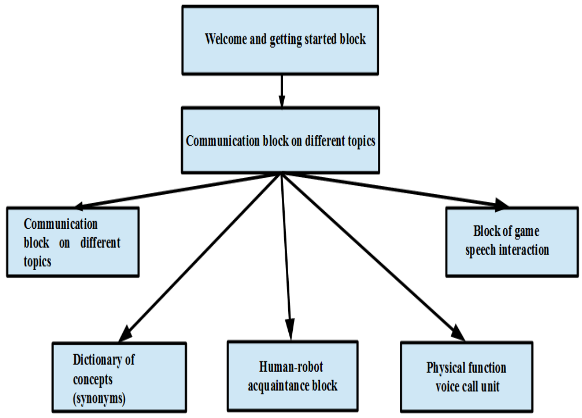

Fig. 1. Flowchart of the system of natural language communication with a NAO robot .

At the same time, for each recognized ID number, a physical action is defined and activated using that number. The semantic network of the game module is shown in Fig. 2.

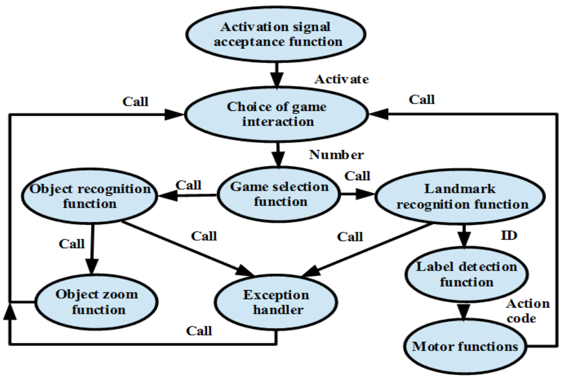

Fig. 2. Semantic network of the game module .

In the human-robot familiarity block, the function of detecting the presence of a person in the robot's field of view is activated. The robot starts to track the position of people, if they appear in its field of view, and tries to identify the person by his image, if the images of this person are available in the database. If this person is unknown to the robot, the robot will offer to get acquainted through a dialog and in the process associate the image of the face with the person's ID and enter this information into the database. If the person wants to become unfamiliar to the ro If the person in question is unknown to the robot, the robot will offer to get acquainted through the introduction dialogue and will link the face image and the person's ID in the process, entering this information into the database. If the person wants to become unfamiliar to the robot again, the robot will activate the erase function of the database upon request. The semantic network of the dating module is shown in Fig. 3.

The block of voice call physical functions allows a number of information exchange processes to be controlled during the dialogue. This block can transmit information to send e-mails and send a signal to activate the function of reading incoming e-mails. The multimedia content recording function is called up, allowing the robot to take a photo, record a video or record a voice memo and play it back.

Authorized licensed use limited to: Universidad de Talca. Downloaded on June 27,2025 at 16:04:03 UTC from IEEE Xplore. Restrictions apply.

Audio files in .mp3, .wav and .ogg format can be played back while adjusting the volume level of the NAO robot's speakers.

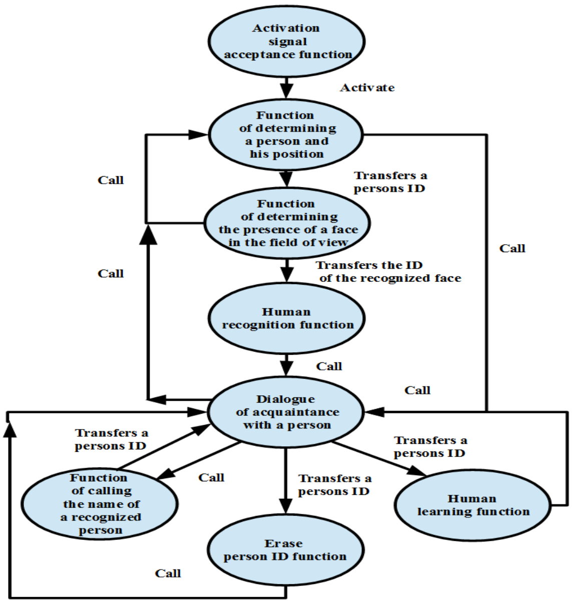

Fig. 3. Semantic network of the dating module .

The robot's infrared transceivers allow for infrared (IR) signal detection, reading of transmitted code and the transmitting of control button codes for TVs and set-top boxes over a distance of several metres. The robot can test the connection to the internet and the Choregraphe development environment, reporting the results to the user. It can generate a pseudo-random number by saying it aloud in a preset natural language. The robot can simulate breathing on demand, perform an action of the interlocutor's choice, and check the temperature of its motors. With the help of visual sensors, NAO can recognise obscured areas in the visible image as well as check their functionality.After each function is performed, a dialogue is returned to, so that the robot does not stop at one function, but provides the interlocutor with the choice of the next action. The semantic network of the functional module is shown in Fig. 4.

III. Behavioral Algoritms Of An Antropomorphic Robot In The Process Of A Natural-Language Dialogue With A Human

Development of the whole project of the robot-human communication was performed in the Choregraphe development environment. For the realization of various physical functions we used blocks with the parameters from the built-in block library. The links themselves, which provide communication, were created with the help of the tool for creating complex dialogs QIChat.The software model of the basic module of the system of natural language communication, by means of which the transition between modules and dialogs is carried out, is presented in Fig. 5.

The Hello! container consists of three blocks: Set Language, which Language parameter value is set to "Russian", Hello, which is a greeting wave animation, and SayHello, in which the robot greets the human. The Say module is responsible for saying the target phrase. The Voice shaping slider is responsible for shaping the voice and Speed for the pronunciation speed. The Set Language module is responsible for setting the language of speech and perception of the robot. In the HELLO module, the main block is the dialogue "Privet, chelovek!". But first, the signal comes to the Basic Awareness block, which determines the presence of a person nearby. If no people are found, the signal goes to the dialogue block. If there are people nearby, then they are determined for their presence in the database. In the absence of them in the database, the robot offers to get to know, that is, to enter there. The blocks that make acquaintance are connected to the dialogue: Name determines the name, Unlearn Face deletes information from the database, Learn Face enters information there and Face Reco. recognizes a person.

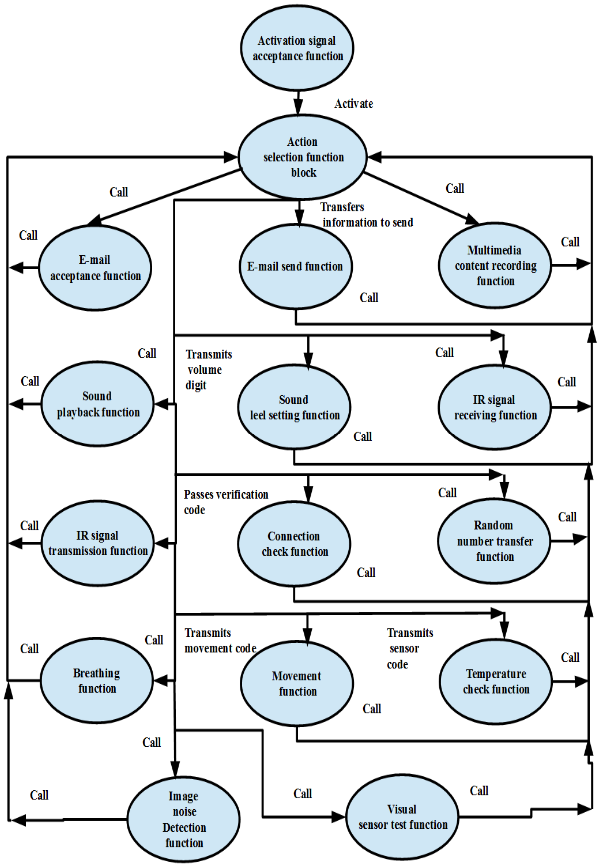

Fig. 4. Semantic network of the functional module.

The RELATIONSHIPS and Concept dialog boxes are written using scripts from the QIChat tool. The Talking container contains three dialogs on the topics of weather, sports, and robots (Fig. 6). The dialogs run simultaneously, transitioning through the central Relationships node, but only one dialog is active, and the focus changes from one topic to another when transitioning.

The tag is denoted by a special symbol “%”. It allows one or more sentences or rules to be identified using a tag. Tags can be activated or deactivated. If the tag is deactivated, then the offer is false and will not be told. In a dialogue about the weather, the conversation begins with the robot's phrase, then he asks the person and, depending on the answer, continues to speak in one of several options. As a result, the conversation boils down to an offer to talk on a different topic, or to start the conversation anew, choosing other answer options as desired.

Authorized licensed use limited to: Universidad de Talca. Downloaded on June 27,2025 at 16:04:03 UTC from IEEE Xplore. Restrictions apply.

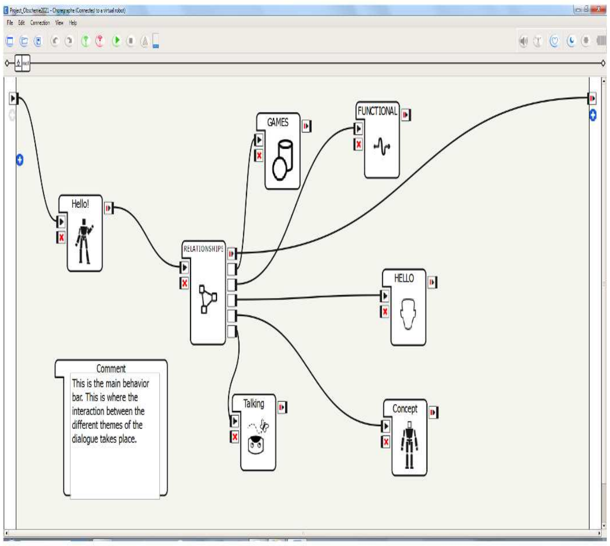

Fig. 5. The main module of the program of communication between the robot and the human.

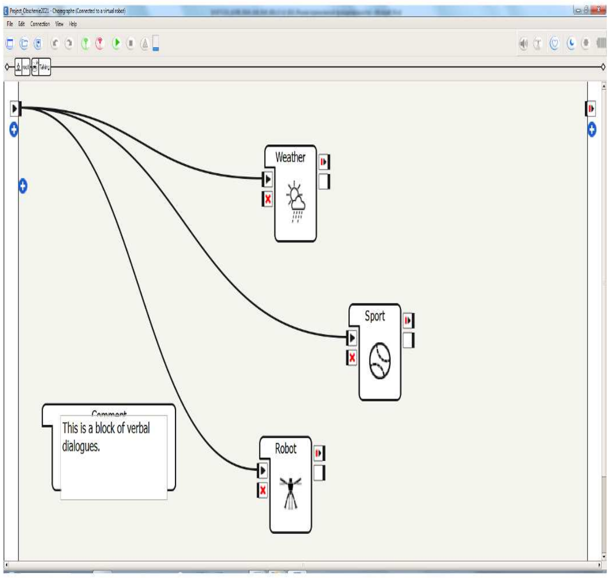

Fig. 6. Talking container .

The essence of the dialogue about sports is that the robot asks what kind of sport the interlocutor likes best, and based on his answer tells information about it. Also, the robot informs about what he himself would like to do.

The essence of the dialogue about robots is that NAO will offer to talk about robots that it knows about itself. He knows about Azimo, R-600, Darwin and Jennifer. After the story, he will assume that the interlocutor will want to know if the NAO itself can perform the same functions as other robots.

By default, the robot's proposal is executed only once, then it is deactivated. One way to reactivate it is to reload all expressions in the theme. To do this, this parameter is applied, while it will only work if the topic is in focus and all its statements are deactivated. Proposal is a robot saying. It is called by a special parameter and can be activated without the interlocutor's phrase. Calling “nextProposal” activates the first active sentence of the robot. Calling “previousProposal” repeats the penultimate pronouncement of this topic.

The call to “sameProposal” repeats the last spoken sentence of this topic. It is recommended that when using it, make the robot say something like "Ok, I will repeat", "no problem" so that the user knows what to expect. It also makes the repetition clearer. If you use the transition to a deactivated dictum by calling “goto”, then nothing will happen. For the phrase to be pronounced, you need to use “gotoReactivate” (tag).

QIChat provides two types of tag navigation, which are called tag. The first allows you to navigate through tags only within one topic, it is called “goto” (tag). The second way is to navigate through tags in any topics, to do this, write “topicTag” (Topic1, tag). The name of the target topic is indicated in brackets of this transition, separated by commas, the name of the tag.

In Choregraphe, QIChat I / O events can have the same name. If the event has the same name as the I / O field, the variable is only associated with the field and is no longer used with ALMemory. It allows you to trigger any NAOqi events that provide information related to the robot: movement, battery information, face information. The event will be executed whenever the topic is activated, but will stop when the topic is deactivated. QIChat has the ability to call the physics animation of the robot from its built-in repository. For this, special commands are used. The “mode” command changes the body language mode for the current text. The “run” command stops the phrase from being spoken, starts the animation, and continues the phrase. The “start” command starts the execution of a behavior or animation. The “stop” command stops the execution of a behavior or animation. The “wait” command makes the robot wait for the animation or behavior to complete completely.

The program model of the game module is presented in Fig. 7. When a signal is sent to this module, the Games dialogue is activated. It contains three games: recognition of a familiar object, recognition of special marks, and text quest. The largest and most diverse module in the project is FUNCTIONAL, which contains the Interface block control dialog box where the selection and access to many different functions of the robot takes place (Fig. 8).

Using this module, the interlocutor can ask NAO to send an Email, receive it, play a music track, record a voice note and then play it back, set the optimal speaker volume level, check the reception of the IR signal, its code.

Also, the robot can send the code of any button via the infrared port to a television device, check the temperature of its motors, record a video, take a photo, check the light of the image from the visual sensors, perform any movement, imitate breathing, say a random number in the range from one to ten, and check your internet connection and Choregraphe. The signal is transmitted here not only of the "Bang" type, but also in the form of text strings of the String type and numbers of the Integer type. Creating several blocks of dialogue allows you to specify different topics of communication and build branched scripts of dialogues. Communication between these modules is carried out through the script of the main topic.

Authorized licensed use limited to: Universidad de Talca. Downloaded on June 27,2025 at 16:04:03 UTC from IEEE Xplore. Restrictions apply.

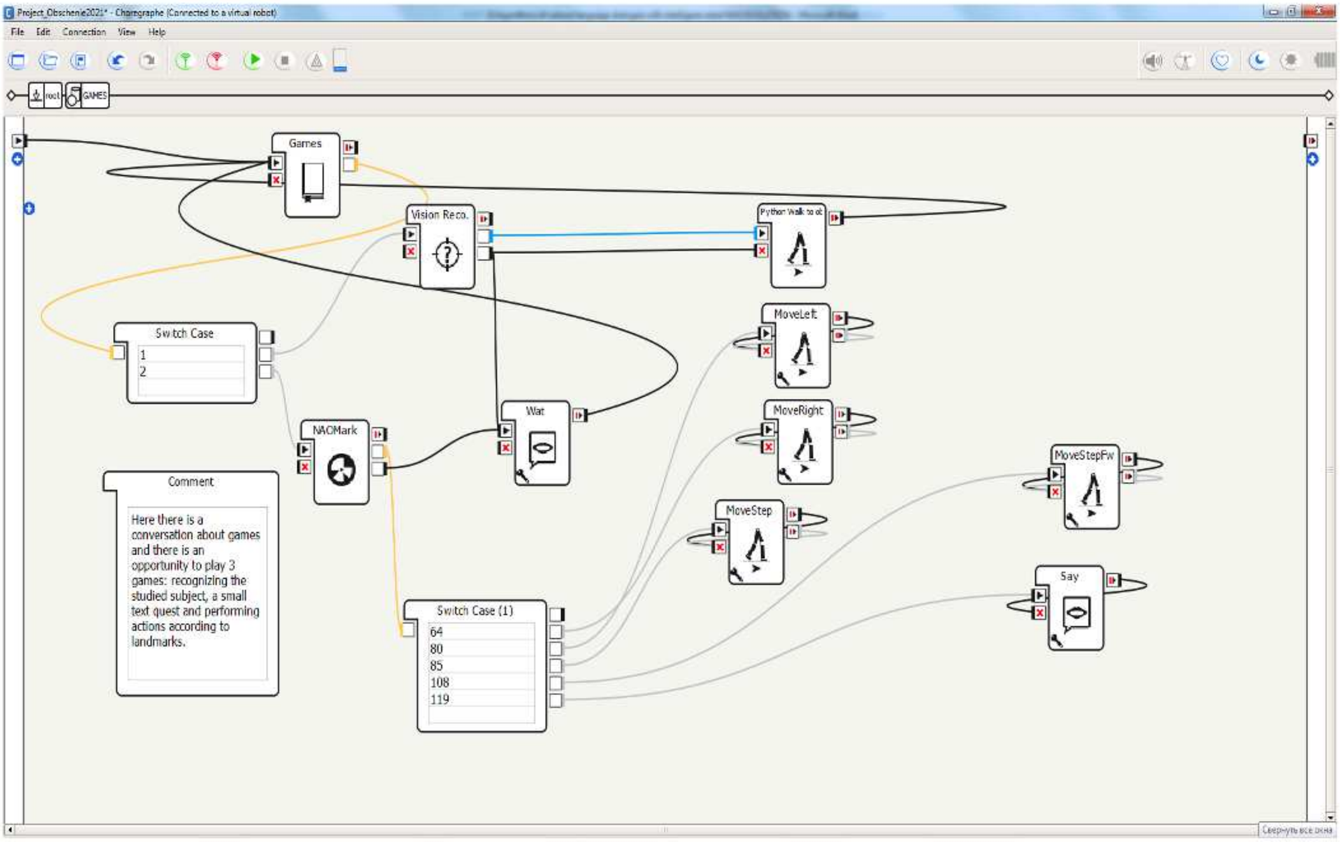

Fig. 7. Game module GAMES.

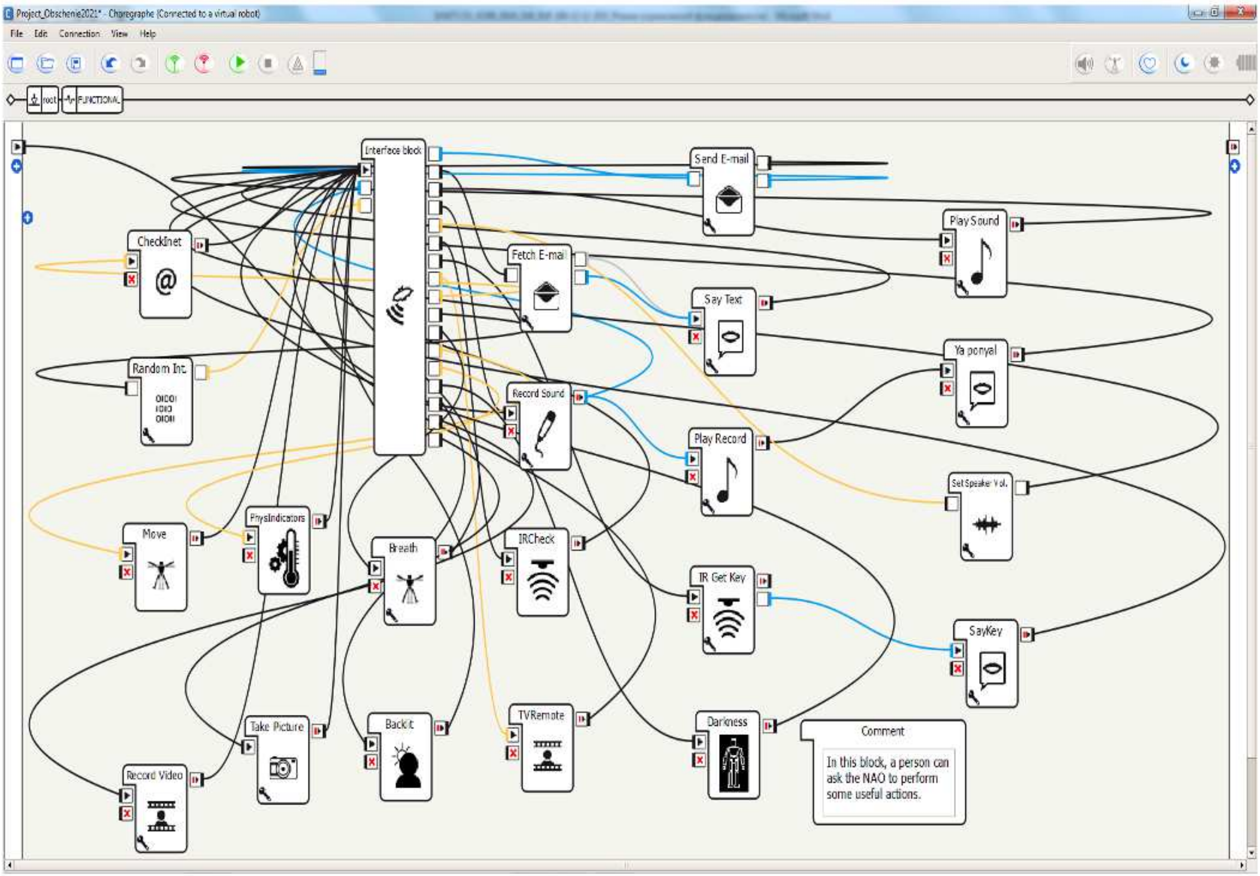

Fig. 8. Function module FUNCTIONAL.

Authorized licensed use limited to: Universidad de Talca. Downloaded on June 27,2025 at 16:04:03 UTC from IEEE Xplore. Restrictions apply.

To do this, the key phrase activating this dialog is indicated first in the topic, and then “private” is entered in each subsequent line. This keyword indicates that these phrases will be active only in this dialogue. Such expressions will not be recognized if another topic is active.

The transition between the topics of the conversation is carried out by pronouncing the key phrase, even if another dialogue is already ongoing. One of the scenarios for organizing dialogue in natural Russian includes two topics.

The Interfaces dialog is activated by the “%” funct tag. The robot invites the interlocutor to find out what he can ask him to do. Since there are many functions, it is divided into two phrases, if a person does not get tired of listening, then he can confirm further listening. The first function is to send an e-mail to the desired address, lines are transmitted sequentially to fill in the parameters for sending the Send Email block. The second function is to read the last email using the Fetch Email block. The Say Text block receives the text of the letter and speaks it.

The third function is to play a sound file in the formats .mp3, .wav and .ogg using the Play Sound block.The fourth function is to record a voice note to a file in the robot's memory using the Record Sound block. Then the note is repeated in the Play Record block. The fifth function is to set the speaker volume using the Set Speaker Vol.

The sixth and seventh functions are to check the IR signal and its code using the IRCheck container and the IR Get Key block. The IRCheck container consists of a block for determining the side of receiving the IR Side signal, which passes the name of the side to the voice modules that sound it. The eighth function is to control television devices using the TVRemote container.

This container consists of the Switch Case selection block, to which the number of the button to be pressed is transferred from the dialog. To find out which number is suitable for which button, the user needs to study the manual presented in the sixth chapter. Each IR Sent Key block has its own TV remote control button, it transmits its code to the device.

The ninth function is to check the Internet connection and Cho-regraphe using the CheckInet container. The tenth function is to simulate the breathing of the robot using the Breath block. The parameters of this block set the heart rate of the robot and the amplitude. The eleventh function is to perform various actions using the Move container. In the Move container, a movement code is received from the dialog, which activates the required movement using the Switch Case selection block. The Stand Up block makes the robot stand up, Move To walk a step, Rest sit down to rest, WakeUp to wake up, Sit Down to sit down.

The twelfth function is to test different sensors of the robot using the PhysIndicators container. The dialog sends the check number, which activates the required check through the Switch Case. Battery checks the battery charge of the robot, Sonar for the presence of obstacles in front of it, Temperature head temperature, Heat heating of motors, TactHead touching the head, TactLHand touching the left hand, TactRHead touching the right hand.

The thirteenth function is to check the robot's visual sensors using the Darkness container. In this container, the presence of a signal from the visual sensors is first checked, and then the result is reported. The fourteenth function is to check for image highlights using the Blacklit container. In this container, the light in the image is checked and it is reported whether it is there, not, or the robot is not sure. The fifteenth and sixteenth functions are video and photo recording using the Record Video and Take Picture blocks.

The parameters of these blocks specify the quality of the survey and the resolution of the output file, as well as the format. You can also select the camera to shoot.

The work on linking different topics of conversation with the transitions from one dialogue to another, using the QiChat dialogue creation tool, is currently underway. It allows you to define different rules of behaviour, manage transitions between conversation topics, use concepts to build a knowledge base of concepts and relations, apply focus to refine the semantics of conversations, and provide many other features.

Especially interesting is the possibility of connecting a dialogue engine which can not only generate robot responses itself, but also connect to the databases of Aldebaran Robotics. This will make it possible to use the existing dialogue database and algorithms for recognising and processing user requests. Further development of natural language communication methods involves extending the QiChat dialogue module by writing scripts in the Python programming language.

CONCLUSION

Use of intelligent anthropomorphic robots in the process of training bachelors and masters in "Information systems and technologies" increases the motivation of students and builds teamwork skills in the implementation of complex projects. Research, development and testing of modules for verbal interaction between the student and anthropomorphic robot provides practical use of theoretical knowledge in the course of laboratory works, opens up opportunities to expand topics and spheres of potential application of the results of graduate qualification works.

REFERENCES

[1] University 2035: Annual Report 2020 https://reports.2035.university/annual_2020 (in Russian)

[2] https://ump.vogu35.ru/index.php/vse-materialy/category/45-informatsionnye-sistemy-i-tekhnologii (in Russian)

[3] Aldebaran Robotics official website: https://www.ald.softbankrobotics.com/en

[4] Robotics in Russia: educational landscape. Part 1 / D. A. Gagarin, A. S. Gagarin; National Research University Higher School of Economics, Institute of Education. – M.: HSE, 2019. – 108 p. – (Modern education analytics. No. 6 (27)). (in Russian)

[5] UNESCO Report on Robotics. - URL: http://unesdoc.unesco.org/images/0025/002539/253952E.pdf (accessed 18.11.2021). - Text : electronic.

[6] Intelligent Robotics in the Structure of the Direction Information Systems and Technologies / A. Shvetcov, V. Gorbunov, K. Kinyakin, S. Dianov // 2020 5th International Conference on Information Technologies in Engineering Education, Inforino 2020 - Proceedings : 5, Moscow, April 14-17, 2020. - Moscow, 2020. - P. 9111807. - DOI 10.1109/Inforino48376.2020.9111807.

[7] Mironov A.A., Kinyakin K.S., Shvetsov A.N. Application of the NAO Anthropomorphic Robot for Solving Intellectual Problems / University Science – Region // Proceedings of the XV All-Russian Scientific Conference with international participation. – Vologda: Publishing House of the VogU, 2017. – pp. 82-83. (in Russian)

[8] Choreography website: https://developer.softbankrobotics.com/us-en/downloads/nao-v5

Authorized licensed use limited to: Universidad de Talca. Downloaded on June 27,2025 at 16:04:03 UTC from IEEE Xplore. Restrictions apply.