THE BASIC ELEMENTS FOR COGNITIVE MODEL OF SPEECH PERCEPTION MECHANISM ON THE BASE OF MULTI-AGENT RECURSIVE INTELLECT

Z.V. NAGOEV[1], I.A. GURTUEVA[2]

[1] Federal state budgetary scientific establishment "Federal scientific center
"Kabardin-Balkar Scientific Center of the Russian Academy of Sciences"
360002, KBR, Nalchik, 2, Balkarov street
E-mail: cgrkbncran@bk.ru
[2] Institute of Computer Science and Problems of Regional Management –
branch of Federal public budgetary scientific establishment "Federal scientific center
"Kabardin-Balkar Scientific Center of the Russian Academy of Sciences"
360000, KBR, Nalchik, 37-a, I. Armand St.
E-mail: iipru@rambler.ru

In this paper, the generalized architecture used in almost all modern systems of automatic speech recognition is analyzed. The necessity of developing a fundamentally new approach to solving speech recognition problems is outlined. A formal description of the structure of the speech perception act is proposed for use as a general theoretical basis in the development of universal automatic speech recognition systems that are highly effective in conditions of high noise and "cocktail party" situations. The general structural dynamics of the speech recognition process has been developed, which allows to take into account the linguistic and extra-linguistic aspects of a speech message. The concept of an articulation event as a minimal basic pattern of sound image recognition has been proposed. The recognition process is structured based on the functional determinants of the situation. The need to analyze the numerous sources of information accompanying the sound message, the rejection of the search for an invariant here is of fundamental nature. Multi-agent systems were chosen as the formal means for implementation. Multi-agent approach allows to differentiate and analyze sounds of different nature. This makes the proposed model unique and gives it advantages in the so-called "cocktail party" situation, as well as in tasks where the noise level is extremely high.

Keywords: artificial intellect, multi-agent systems, speech recognition, artificial neural networks.

REFERENCES

1. Abdel-Hamid O., Mohamed A., Jiang H., Penn G. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition, Proc. IEEE Int. Conf. Acoust., Speech, Signal Process, 2012. Pp. 4277-4280.

2. Chomsky N.A. A Review of Skinner's Verbal Behavior, [Readings in the Psychology of Language]. Prentice-Hall, Upper Saddle River, New Jersey, 1967. Pp. 636.

3. Coates A., Ng A.Y. Learning feature representations with *k*-means. *Neural Networks: Tricks of the Trade,* 2012. Pp. 561-580.

4. De Mulder W., Bethard S., Moens M.-F. A Survey on the Application of Recurrent Neural Networks to Statistical Language Modeling // Computer Speech and Language. 2015. №30(1). Pp. 61-98.

5. De Saussure F. *Kurs obshei lingvistiki* [Course in General Linguistics]. *Yekaterinburg: Izdatel'stvo Ural'skogo Universiteta*, 1999. Pp. 256.

6. Deng L., Li X. Machine Learning Paradigms for Speech Recognition: An Overview, IEEE Transactions on Audio, Speech, and Language Processing, 2013. № 21(5). Pp. 1060-1089.

7. Gazzaniga M.S. Conversations in the Cognitive Neuroscience. The MIT Press, Cambridge, 1996. Pp. 752.

8. Ghai W., Singh N. Literature Review on Automatic Speech Recognition // International Journal of Computer Applications. 2012. № 41 (8). Pp. 42-50.

9. Ghitza O. Auditory nerve representation as a front-end for speech recognition in a noisy environment // Computer Speech and Language, 1986. Vol. 1. Pp. 109-130.

10. Gupta V. A Survey of Natural Language Processing Techniques // International Journal of Computer Science & Engineering Technology. 2014. № 5 (1). Pp. 14-16.

11. Haikonen P. The Cognitive Approach to Conscious Machines, imprint Academic, Exeter, UK, 2003. Pp.300.

12. Hinton G., Deng L., Yu D., et al Deep neural networks for acoustic modeling in speech recognition, IEEE Signal Process. Mag, 2012. № 29(6). Pp. 82-97.

13. Juan B.H. Speech Recognition in Adverse Environments // Computer Speech and Language, 1991. Vol. 5. Pp. 275-294.

14. Jurafsky D., Martin J. Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition, Prentice Hall, Boston, 2008. Pp. 1032.

15. Kotseruba I., Tsotsos J.K. A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications, 2016 available at: arxiv.org/abs/1610.08602.

16. Mazurenko I.L. *Komp'utrenye sistemy raspoznavaniya rechi* [Computer Speech Recognition Systems] // *Intellektual'nye sistemy* [Intellectual Systems], 1998. Vol. 3. issue. 1-2. Pp. 117-134.

17. Minsky M. The Society of Mind. Simon and Shuster. New York, 1988. Pp. 336.

18. Mohamed A., Dahl G., Hinton G. Acoustic modeling using deep belief networks, IEEE Audio, Speech, Lang. Process, 2012. № 20(1). Pp. 14-22.

19. Morozov V.P., Vartanyan I.A., Galunov V.I. *Vospriyatiye Rechi: voprosy funktsional'noi asimmetrii mozga* [Speech Perception: Issues of functional brain asymmetry]. *Leningrad: Nauka,* 1988, 135 c.

20. Nagoev Z.V. *Intellektika, ili Myshlenie v zhivych i iskusstvennych sistemach* [Intellectics or Thinking in Living and Artificial Systems]. *Nalchik: Izdatel'stvo KBNC RAN*, 2013. Pp. 232.

21. Nagoev Z.V., Nagoeva O.V. *Izvlechenie znanii iz mnogomodal'nyh potokov nestrukturirovannyh dannyh na osnove samoorganizatsii mul'tiagentnoi kognitivnoi arhitektury mobil'nogo robota* [Knowledge Extraction from Multimodal Streams of Unstructured Data on the Base of Self-Organization of Multi-Agent Cognitive Architecture for Mobile Robot], *Izvestia KBNC RAN* [News of KBSC of RAS]. 2015. № 6 (68). Pp. 73-85.

22. Nagoev Z.V., Nagoeva O.V. *Zritel'nyi analizator intellektual'nogo robota dlya obrabotki nestrukturirovannyh dannyh na osnove mul'tiagentnoi neirocognitivnoi arhitektury* [Visual Analyzer of Intellectual Robot for Unstructured Data Processing on the Base of Multi-agent Neurocognitive Architechture] // *Perspektivnye sistemy i zadachi upravleniya: Materialy vserossiiskoi nauchno-prakticheskoi konferencii* [Advanced Systems and Management Tasks: Proceedings of the 12th All-Russia Conference]. Rostov-on-Don, 2017. Pp. 457-467.

23. Nagoev Z.V., Denisenko V.A., Lyutikova L.A. *Sistema obucheniya avtonomnogo sel'skohozyaistvennogo robota raspoznavaniyu staticheskih izobrazhenii na osnove multiagentnyh kognitivnyh arhitektur* [Learning System of Autonomous Agricultural Robot for Static Images Recognition on the Base of Multi-Agent Cognitive Architectures] // *Ustoichivoie razvitie gornyh territorii* [Sustainable Development of Mountain Territories]. 2018. № 2. Pp. 289-297.

24. *Nagoev Z., Lyutikova L., Gurtueva I.* Model for Automatic Speech Recognition Using Multi-Agent Recursive Cognitive Architecture, Annual International Conference on Biologically Inspired Cognitive Architectures BICA, 2018, Prague, Chech Republic http://doi.org/ 10.1016/j.procs.2018.11.089

25. Newell A. Unified Theories of Cognition, Harvard University Press, Cambridge, Massachusetts, 1990. Pp. 576.

26. Rabiner L.R., Schafer R.W. *Tsifrovaya Obrabotka Rechevyh Signalov* [Digital Processing of Speech Signals]. Moskva: Radio I Svyaz', 1981. Pp. 496.

27. Reddy R. Speech Recognition by Machine: A Review, Proceedings of the IEEE, 1976. № 64 (4). Pp. 501-531.

28. Ronzhin A.L., Karpov A.A., Kagirov I.A. *Osobennosti distantsionnoi zapisi i obrabotki rechi v avtomatah samoobsluzhivaniya* [Peculiarities of Remote Recording and Speech Processing in Self-Service Machines]. // *Informatsionno-upravlyayushie sistemy* [Information and Control Systems], 2009. № 5. Pp. 32-38.

29. Schunk D. H. Learning Theories: An Educational Perspective, Pearson Merrill Prentice Hall, Boston, 2011. Pp. 576.

30. Van Veen B.D., Buckley K.M. Beamforming: A Versatile Approach to Spatial Filtering, IEEE ASSP Magazine, 1988. № 5(2). Pp. 4-24.

31. Waibel A., Lee K.-F. Readings in Speech Recognition, Morgan Kaufman, Berlington, 1990. Pp. 680.

32. Wooldridge M. An Introduction to Multi-Agent Systems, Wiley, Hoboken, 2009. Pp. 366.

33. Zinder L.R. *Obshaya Fonetika* [General Phonetics]. *Moscow: Vysshaya Shkola*, 1979. Pp. 312.

34. Zion Golumbic E.M., Ding N., Bickel S., et al. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party", 2013. Neuron, 77 (5). Pp. 980-991.

Nagoev Zalimkhan Vyacheslavovich, Candidate of technical sciences, acting Chairman of Kabardin-Balkar Scientific Center of the Russian Academy of Sciences.

360000, KBR, Nalchik, I. Armand street, 37-a.

Ph.: (8662) 42-65-62, fax: (8662) 42-65-62.

E-mail: zaliman@mail.ru

Gurtueva Irina Aslanbekovna, staff scientist of the Department of computer linguistics, Institute for Computer Science and Problems of Regional Management, Kabardin-Balkar Scientific Center of RAS.

360000, KBR, Nalchik, I. Armand street, 37-a.

Ph.: (8662) 42-65-62, fax: (8662) 42-65-62.

E-mail: gurtueva-i@yandex.ru