# Self-learning Semantic-distance-based Answering System with Automatic Morpheme Recognition

L. Dudás
Department of Information Engineering, University of Miskolc, Miskolc, Hungary

***Abstract*—At the beginning of the 21st century humanity has accumulated a very large amount of static data and facts. The problem is to make it active, to produce knowledge. One of the keys to this is restructuring, representing data in a new way. The emphasis is on the dynamic associations between data elements. Instead of storing distinct atoms of knowledge the goal is to build up a network of facts and methods that mirrors the relations found among them in the world, moreover in the human brain. One of the difficult tasks in this endeavor is to capture the meaning. The most feasible knowledge representation method for this purpose is the semantic network. This powerful technique has its new Renaissance in our days. This article offers a new definition of the meaning of a sentence for a human and presents a simple technique for searching for sentences having meanings close to the meaning of a given sentence. In some languages the morpheme structure of words plays a particularly important role in this process. This paper presents an automatic morpheme learning technique integrated in the answering system, using the Hungarian language as an example.**

***Keywords:*** **NLP, self-learning, semantic distance, answering system, meaning, morpheme recognition.**

## REFERENCES

[1] S. Abney, M. Collins and A. Singhal, "Answer extraction", *Proceedings of ANLP 2000.* Seattle, pp. 296–301, April 2000.

[2] Z. Zheng, "Developing a web-based question answering system", http://www2002.org/CDROM/poster/203/, Accessed: 2 September 2011.

[3] J. D. Litman and P. Shimei, "Designing and evaluating an adaptive spoken dialogue system", *User Modeling and User-Adapted Interaction*, vol. 12(2/3), pp. 111–137, 2002.

[4] S. Small, T. Strzalkowski, T. Liu and N. Shimizu, "HITIQA: an interactive question answering system – a preliminary report", *Proceedings of ACL '03, Workshop on QA*, Sapporo, Japan, 2003.

[5] C. O' Riordan and H. Sorensen, "Information filtering and retrieval: an overview", http://scholar.google.hu/scholar?q=%22 Information+filtering+and+retrieval:+an+overview%22&hl=hu&as_sdt=0&as_vis=1&oi=scholart, Accessed: 2 September 2011.

[6] L. Kovacs, "A fast algorithm for building concept set", *Proceedings of MicroCAD 2002 International Scientific Conference*, Miskolc, Section H, p.113-118, March 2002.

[7] Customized Computing – AI links, http://spydaz.tech.officelive.com/chatbots.aspx, Accessed: 3 September 2011.

[8] D. Davidson, *Inquiries into Truth and Meaning*, 2nd ed., Oxford, 2001.

[9] *Artificial Intelligence (Ch 5) Knowledge-Based Agents*, http://physicsarchives.com/index.php/courses/243, Accessed: 2 September 2011.

[10] S. Bird, P- Blunsom, "Human language technology - Lexical Semantics and Word Sense Disambiguation", http://ww2.cs.mu.oz.au/460/2004/materials/wk8.pdf, Accessed: 3 September 2011.

[11] J. J. Katz, *Propositional Structure and Illocutionary Force – A Study of the contribution of sentence meaning to speech acts*, Harvard University Press, 1980.

[12] D. Warren and P. Fernando, "An efficient easily adaptable system for interpreting natural language queries", *American Journal of Computational Linguistics*, vol. 8, pp. 110–122, November 1982.

[13] C. Callison-Burch and P. Shilane, "A natural language question and answer system", in *Unpublished Manuscript Stanford University Stanford CA*, 2000., http://wenku.baidu.com /view/eaa2a00e4a7302768e99395d.html, Accessed: 12 August 2011.

[14] F. Bodon, "Ragelemző szoftver – Suffix analyser", http://www.cs.bme.hu/~bodon/magyar/program/c++/rag/ragelemzes.htm#nevezetes, Accessed: 3 September 2011.