Three Address Code Based Semantics Processor for Sinhala

IWMHD Bandara[1], B Hettige[#1] and DDM Ranasinghe[2]

[1]*Department of Computer Engineering, General Sir John Kotelawala Defence University, Sri Lanka*
[2]*Department of Electrical and Computer Engineering, The Open University, Sri Lanka*

[#]budditha@kdu.ac.lk

Semantic processing techniques have a wide interest in the field of Natural Language Processing. Processing a semantic from a natural language for human-machine communication is still a research challenge in this field. The Three-Address-Code is a type of intermediate code used by the compilers to identify the meaning of the source code or statements easily, with full accuracy. Therefore, the research captures the semantics of the Sinhala language text through this Three-Address-Code concept. This paper presents a Three-Address-Code based semantic processing system that can be used for human-machine communication using the Sinhala language. The proposed system comprises three components; namely Sinhala Part of Speech tagger, Sinhala chunker, and three-address-code based semantics generator. The system takes the Sinhala sentence as an input and generates the semantics information. This semantic processing system has been used under the PINA system for semantic processing.

Keywords: *three-address-code, semantics, tagging, chunking, ontology*