

SI-642  
SEN

**AGENT APPLICATION FOR WORLD WIDE WEB SEARCHING**

**A PROJECT REPORT PRESENTED BY**

**JANAKA WEERATHUNGA YAPA SENEVIRATNE**

**to the Board of Study in Computer Science of the  
POSTGRADUATE INSTITUTE OF SCIENCE**

*in partial fulfillment of the requirement  
for the award of the degree of*

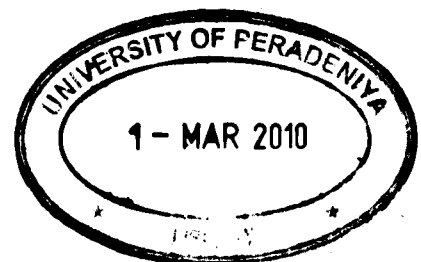
**MASTER OF SCIENCE IN COMPUTER SCIENCE**

**of the**

**UNIVERSITY OF PERADENIYA**

**SRI LANKA**

**2009**



**628211**

# **AGENT APPLICATION FOR WORLD WIDE WEB SEARCHING**

J.W.Y Seneviratne

**Postgraduate Institute of Science**

University of Peradeniya

Peradeniya

Sri Lanka

The objective of this research project is to gain an understanding of natural language processing and web searching technologies with the expectation of applying this knowledge to develop a sophisticated web agent application that can perform context aware web searching. This thesis explains background, requirements, design and implementation details of the proposed web agent.

Natural language processing techniques have been evolved to satisfactory levels thus giving the software applications the commonsense understanding capabilities to a certain degree. Web searching has become so important and common activity for world wide web users. Figuring out ways to improve web searching user experience by using commonsense techniques can leverage the potential usage of web in several ways. Mainly, the view of world wide web as a scattered information and services collection can be fixed to a more organised view by deploying well developed web agent applications in place of current web browsers with limited or no commonsense knowledge capabilities.

The web agent application described here is intended to be used for context aware information search and to find related web services for a particular task thus improving web searching user experience. The web searching agent application developed in this research project can be treated as a step towards the anticipated semantic web, whereby end users can take maximum use of the scattered and differently formatted information on world wide web.