

Deployment of a Natural Language Processing system on Grid

Roshan Sumbaly, Abhishek Kumar, Gaurav Paruthi, Shubham Malhotra
Birla Institute of Technology & Science, Pilani – Goa Campus

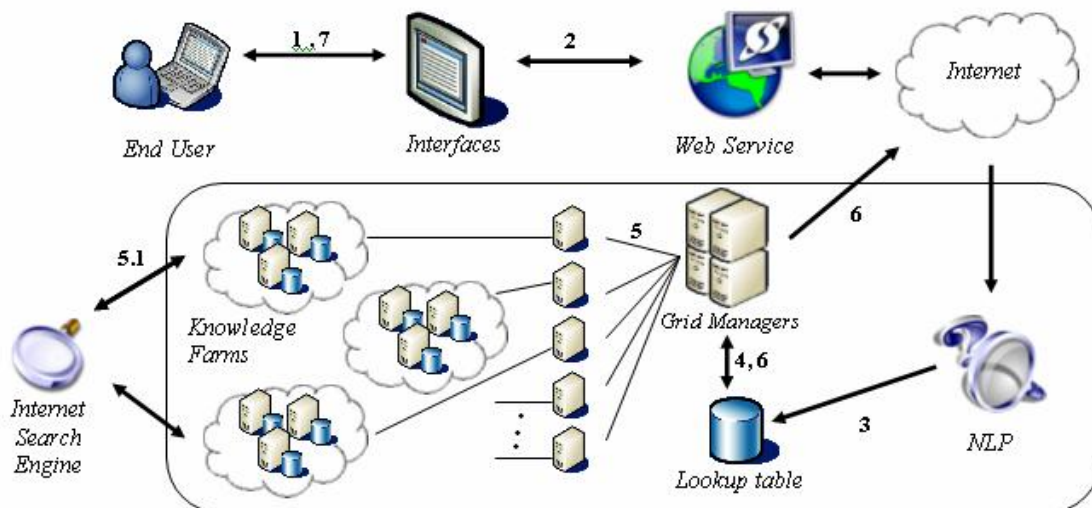
Decades ago knowledge was restricted to certain locations across the globe and this resulted in sporadic knowledge. Now in the 21st century we face a crisis of information overload and hence separating the wheat from chaff is essential today.

Our article presents a Grid based framework, combined with the power of natural language processing (NLP), to act as a virtual assistant. This self-learning virtual assistant can answer queries in an intelligent and conversational manner and is capable of being deployed in various scenarios like an academic setup. Our system has other potential applications such as web-based online marketing and helpdesk diagnostics in different sectors.

We explore the integration of Grid with NLP, an AI subdomain, to produce quick and accurate results. The assistant received queries from the various interfaces using webservice. It then uses NLP to understand the domain of the question. The Grid, then in collaboration with a lookup table, routes the queries to the correct *knowledge farm*, depending on the domain found. We define knowledge farms as a set of nodes in the Grid containing a cluster of information. A second module of NLP, at the knowledge farm level, responds with an accurate conversational answer. Our implementation uses a novel method which involves the working of the Grid and NLP in concert to mine relevant information quickly.

We use Alchemi .NET Grid Framework [4] developed by the University of Melbourne, Australia as our backbone. This Grid Architecture plays the role of a load balancer as it queues up requests before routing; acts as a distributed database by storing and classifying the data across the Grid nodes; and finally, helps in achieving parallelism by accepting queries from multitude users at once. We aim to exploit the robust feature of our Grid to give quick answers to *urgent* queries. Our NLP subsystem uses AIML, i.e. Artificial Intelligent Markup Language, which is an XML dialect for NLP applications.

Our work has received recognition at Imagine Cup [1], a Microsoft sponsored international level competition where we were awarded the Runners-up prize. We will also be presenting a detailed description of our system architecture and implementation at the “International Conference On High Performance Computing” (HiPC 2007) [2] in the form of a poster titled “Artificially Intelligent Grid Assistant (AIGA).”



Overview of our system

Links

- [1] Microsoft Imagine Cup Project Details :

http://www.microsoft.com/india/student/ic07/T_eduGRID.aspx

- [2] Poster titled “Artificially Intelligent Grid Assistant (AIGA)” at International Conference On High Performance Computing (HiPC 2007) :

<http://hipc.org/hipc2007/hipc2007posters.html>

<http://rsumbaly.googlepages.com/AIGA.pdf>

- [3] Relevant News Clipping

http://economictimes.indiatimes.com/News/News_By_Industry/Infotech/Software/Microsoft_announces_team_of_Indian_student_innovators/articleshow/2004855.cms

http://sandpaper.bitsaa.org/notes_goa/success.htm

http://timesofindia.indiatimes.com/Opinion/Sunday_Specials/Innovative_ideas_Look_to_college_kids/rssarticleshow/2076728.cms

<http://www.openpr.com/news/20147/MicrosoftImagineCup2007BITSPilaniGoaCampusInnovatorswinAcclaim.html>

<http://www.newswiretoday.com/news/17793/>

- [4] Akshay Luther, Rajkumar Buyya, Rajiv Ranjan, and Srikumar Venugopal

Alchemi: A .NET-Based Enterprise Grid Computing System,

Proceedings of the 6th International Conference on Internet Computing (ICOMP'05), June 27-30, 2005, Las Vegas, USA.