Improving the Performance of Existing Information Retrieval Systems Using a Software Agent

MAJ Bernard J. Jansen
Department of Electrical Engineering and Computer Science
United States Military Academy
West Point, New York 10996 USA
1 914 938 3233
jjansen@acm.org

EXTENDED ABSTRACT

This paper describes a software agent developed specifically for integration with existing information retrieval interfaces and search engines. The software agent assists the user with query reformulation. The agent assistance is based on characteristics of the user population, user actions during the search process, information from retrieved documents, and historical information from past queries. With minor modification, the software agent can be integrated with a variety of interfaces and search engines.

Keywords

Software agents, information retrieval, adaptive interfaces, interface agents, software integration.

See Other Publications

INTRODUCTION

The amount of information available via networks and databases has rapidly increased and continues to increase. Most existing information retrieval (IR) systems provide limited assistance to users in locating the relevant information that they need. Much research has focused on designing entirely new IR systems. Given the development costs, organizational friction, and system transfer expenses, the total replacement approach will probably have limited impact. This research project is exploring the feasibility of combining software agents with effective IR techniques to improve the performance of existing IR systems. The resulting package will work with any interface or search engine with minimal alternation to the communication channels. Other researchers have explored this idea of application and agent integration [2].

Software Agent Package

Focusing on users of the Web as a diverse population, I draw on a large analysis of the user population [1] to develop user characteristics. Using these characteristics and other information, the software agent assists the user in query formulation. The query is the key element in translating the user's information need to a format the search engine can act upon.

In this prototype, the interface agent draws on characteristics derived from the user population analysis, monitors user actions during the search process, and gleans information from previous query results. Using this collected information, the agent autonomously checks and corrects certain common query formulation

mistakes, suggests possible query terms to the user based on retrieved documents and a thesaurus, and recommends terms to remove from the query. The prototype currently uses a "typical" looking Web browser interface and the mg search engine; however, with minimal modification it could work with any interface or search engine.

CONCLUSION

By developing packages for integration into existing IR applications rather than totally rebuilding systems, one can improve the performance of existing applications while avoiding the cost of product replacement. Using known user characteristics, user actions, and IR techniques, a software agent can provide active and beneficial assistance to a user for query reformulation during a search session. Once additional functionality is added to this prototype, I will conduct user testing to measure the usability and performance impact of the software agent.

REFERENCES

1. Jansen, Major Bernard J., Amanda Spink, Tefko Saracevic. (1998) *Real life information retrieval: A Study of User Queries on the Web*. SIGIR Forum, vol. 32, no. 1, P. 5-17
2. Lieberman, Henry. Integrating User Interface Agents With Convention Applications. Proceedings of the 1998 International Conference on Intelligent Interfaces. P.39-46.