

실시간 상호커뮤니케이션에 의한 인력기반 질의응답시스템 : rPodo

임희석* 홍성훈** 류기곤*
한신대학교 컴퓨터정보소프트웨어학부* (주)에이치앤비솔루션**
{limhs, gon0121}@hs.ac.kr* honglee2@hnbs.co.kr**

Human Powered Question and Answering System by using Real-Time Interactive Communication : rPodo System

Heuiseok Lim* Sunghoon Hong** Kigon Ryu*
Computer Information and Software Department
Hanshin University *
Human & Bridge Solution CO., Ltd**

요 약

우리가 현재 사용하고 있는 정보검색 시스템은 사용자의 질의와 연관있는 문서 집합만을 제공하므로 사용자가 원하p23_final_55.pdf는 정답을 찾기 위해서 사용자는 문서 집합을 브라우징하는 수고를 하여야 하며, 이러한 수고를 덜어주기 위해 개발되는 자동 질의응답시스템은 의미 분석 및 지식 추출 등의 기술적 한계로 사용자에게 만족할 만한 서비스를 제공하고 있지 못한 실정이다. 본 논문은 인터넷에 연결되어 있는 사람 중에 질의어에 대한 응답을 할 수 있는 지혜 제공자를 자동으로 검색 분류하여, 질의자와 실시간으로 연결하여 사용자와 지혜 제공자가 실시간 상호커뮤니케이션을 이용하여 지혜를 교류할 수 있는 인력기반 질의응답시스템인 알포도 시스템을 제안한다. 알포도 시스템은 질의응답모듈, 메티스 관리 모듈, 실시간 커뮤니케이션 모듈, 그리고 지식 추출 및 관리 모듈로 구성되며 현재 베타 서비스를 실시 중이다.

1절 서론

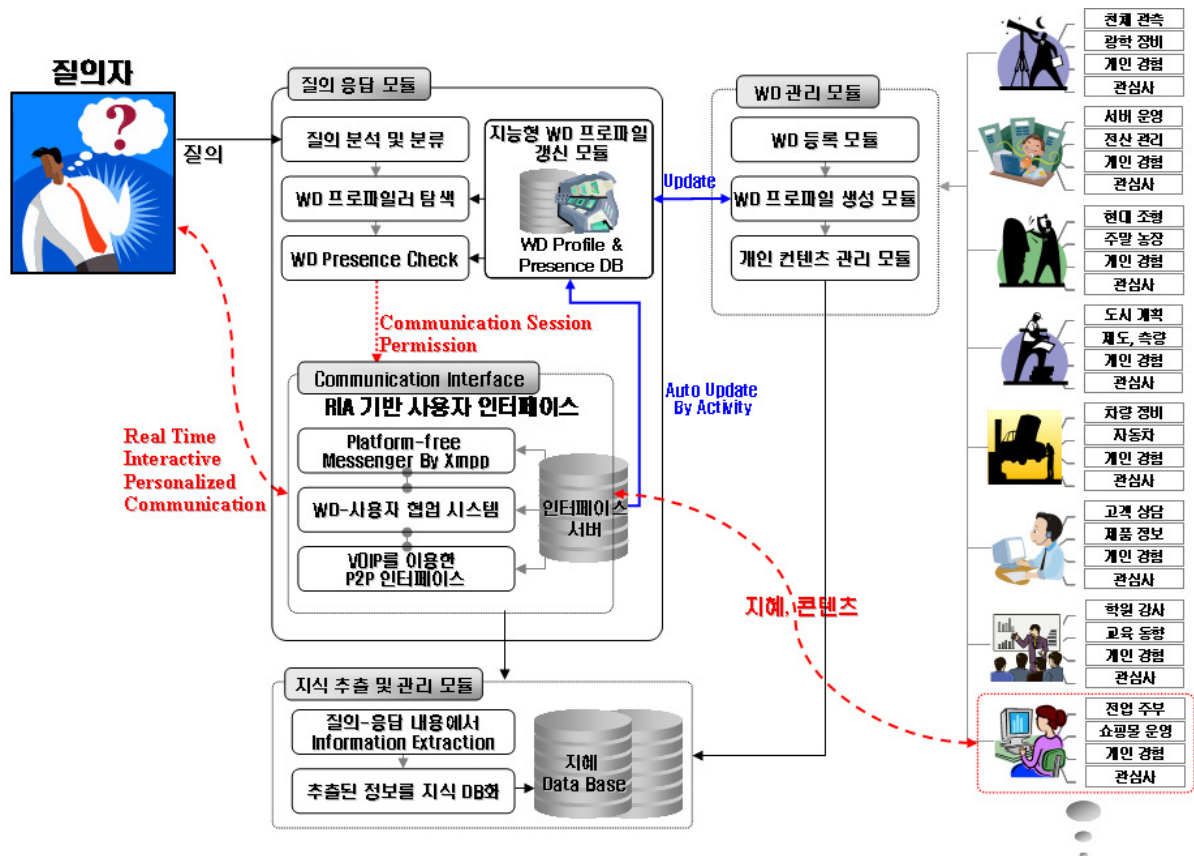
정보검색 시스템은 남녀노소 상관없이 정보를 얻기 위해서 사용하는 매우 유용한 도구가 되었다. 사용자들은 하루에도 몇 번 아니 수십 번 씩 본인이 원하는 지식과 정보를 찾기 위하여 정보검색 시스템을 사용한다. 지난 몇 십년간의 정보 검색 기술의 비약적인 발전과 인터넷 공간의 대용량의 정보 축적으로 인하여 노력하면 본인이 원하는 정보를 빠른 시간에 얻을 수 있는 세상이다.

정보 및 지식 탐색에 있어서 만능 해결사와 같은 정보검색 시스템은 아직 사용자들이 원하는 수준의 정보 서비스를 제공하지는 못하고 있는 실정이다. 정보 검색 시스템은 정보의 수요(information need)가 있는 사용자가 자신의 정보의 수요를 나타낼 수 있는 몇 가지의 단어를 이용하여 질의어에 대한 수없이 많은 연관된 문서(relevant document) 집합을 출력한다. 따라서 사용자는 본인이 원하는 정보나 지식을 얻기 위해서 검색 시스템이 출력한 문서 집합의 문서들을 일일이 브라우징하면서 각 문서를 읽을 수밖에 없다. 하지만 진정으로 사용자들이 원하는 것은 질의어와 연관된 문서의 집합이 아니

라 질의어에 대한 정답(answer)이다. 이러한 사용자의 요구에 부응하기 위한 시스템이 자동 질의응답 시스템(automatic question and answering system)이다. 질의응답시스템은 사용자의 질의에 상응하는 응답을 생성하는 시스템으로 양질의 정답을 생성하기 위해서는 고성능의 질의 의미 분석, 문서의 의미 분석 및 문서에서의 정보추출(information extraction) 등이 요구된다. 하지만 아직까지의 자연어처리 기술과 관련 기술의 한계로 사용자가 만족할만한 질의응답시스템의 개발은 미흡한 실정이다.

기존의 정보검색 시스템과 자동 질의응답시스템은 사용자가 원하는 정답을 생성하기 어렵다는 한계뿐만 아니라 상호성(interactivity), 실시간성(real-time)에서도 한계를 가지고 있다. 기존의 정보검색 시스템과 자동 질의응답시스템은 질의에 대한 단방향의 정보 제공만을 수행할 뿐 사용자의 관심사에 대한 연속적인 질의에 대한 응답을 제공하기 어렵다.

본 연구는 정보의 수요가 있는 사용자와 정보 제공자가 상호 커뮤니케이션을 통하여 자신이 원하는 지혜를 교류할 수 있는 실시간 커뮤니케이션에 의한 인력기반



[그림 1] 알포도 시스템 구성

질의응답 시스템인 알포도(rPodo, <http://rpodo.com>) 시스템을 제안한다. 본 연구에서 지혜란 지식, 정보를 포함하여 콘텐츠, 경험, 노하우, 아이디어 등 타인과 교류할 수 있는 활용가치가 있는 총체적인 것을 의미한다. 전통적으로 사람들의 지혜 획득 방법은 관련 지혜를 가진 사람을 찾아 그에게 의뢰하고 서로 상의하는 방식이었다. 알포도 시스템의 기본 철학은 각기 다른 분야의 지혜와 관심을 가지고 있는 많은 사람들이 항상 연결되어 있는 인터넷에서 지혜를 가진 사람을 찾아 지혜의 수요가 있는 사용자와 연결하여 실시간으로 지혜를 상호 교류하게 하는 것이다. 사람이야말로 그 어떤 시스템보다 지혜 제공에 있어서 지능적이고 효율적이라 할 수 있기 때문이다.

2절 알포도 시스템 개요

알포도 시스템은 web2.0 시대의 키워드인 “개방”, “공유”, “참여” 정신을 가진 온라인상의 지혜를 가진 사람(메티스:metis)를 정확하게 찾아내고 메티스와 질의자가 실시간으로 상호 커뮤니케이션하면서 지혜를 유통, 배포

할 수 있는 웹 2.0 기반의 지혜상호교류시스템으로 [그림 1]은 알포도 시스템의 개요를 도식화한 것이며 알포도 시스템은 크게 4가지의 세부 모듈로 구성된다.

- **질의응답모듈** : 질의 분석, 질의 자동 분류, 그리고 메티스 검색 기능 수행
- **메티스 관리 모듈** : 메티스 등록, 프로파일 자동 갱신 및 자동 분류 기능 수행
- **실시간 커뮤니케이션 모듈** : 질의자와 메티스가 실시간으로 상호 교류하기 위해서 필요한 커뮤니케이션 기능 수행
- **지식 추출 및 관리 모듈** : 지혜 검색 및 자동 질의응답시스템 개발에 활용하기 위하여 질의자와 메티스가 교류한 지혜를 저장 및 가공

3절 질의응답모듈

질의응답모듈은 사용자가 입력한 질의를 분석하고 질의어와 연관이 높은 메티스를 검색하고, 질의를 자동으로 분류하여 지혜를 제공할 수 있는 분야의 메티스를 탐

색하는 기능을 수행한다.

메티스 검색과 질의 자동 분류에 의한 후보 메티스의 목록, 그리고 메티스의 활동 정보 등을 종합하여 사용자에게 출력할 순서화된 메티스 목록을 작성하는 메티스 순서화(metis ranking)도 질의응답모듈에서 수행된다.

4절 메티스 관리 모듈

메티스 관리 모듈은 메티스 등록, 메티스 자동 분류, 그리고 메티스 프로파일의 자동 갱신과 같은 기능을 수행한다. 메티스는 시스템에 등록될 때 본인이 타인에게 전달할 수 있는 지혜 또는 관심사에 대한 내용을 자연어로 기술한다. 메티스 등록기는 이러한 메티스의 정보를 가공하여 메티스 프로파일(metis profile)로 생성한다. 메티스 프로파일은 사용자의 질의와 연관이 있는 메티스를 탐색하기 위하여 사용된다. 메티스 프로파일은 형태소 분석을 통한 명사 추출과 정보 검색의 색인 기법을 이용하여 생성된다. 메티스 자동 분류기는 메티스들의 관심 분야 및 전문 분야를 자동으로 분류하기 위한 모듈이다. 메티스들의 전문 분야는 메티스 등록 시 사용자가 선택을 하지만, 자신의 관심 분야와 전문 분야를 시스템이 제시하는 분류 목록에서 제대로 찾기란 쉽지 않은 일이다. 메티스 자동 분류기는 메티스들이 질의자와 커뮤니케이션을 수행할 때 사용하는 대화 내용을 분석하여 해당 메티스의 관심 분야 및 전문 분야를 자동으로 분류하는 기능을 수행한다.

메티스 프로파일 자동 갱신기는 대화 내용을 분석하여 메티스의 프로파일에 관심 키워드를 추가 및 변경하고, 메티스의 분야를 자동으로 변경하는 기능을 수행한다. [그림 2]는 알포도 시스템에서의 지능적으로 수행되는 순환적 메티스 관리 흐름을 도식화 한 것이다.



[그림 2] 지능형 지혜나눔이 프로파일 관리

5절 실시간 커뮤니티케이션 모듈

메티스 클라이언트 모듈, 메신저 모듈, VoIP 모듈, 협업 모듈 등으로 구성된다. 메티스 클라이언트 프로그램은 네트워크상에 연결된 메티스의 상태(presence)를 시스템 측으로 전달하며, 사용자나 시스템이 메티스를 호출하는 경우 메티스에게 사운드/팝업창 도출/자동 화면전환/화면 깜빡임 등의 수단을 통해 그 사실을 알려준다. 지혜 제공 요청을 받은 메티스는 클라이언트 프로그램을 이용하여 가 사용자가 필요로 하는 지혜를 쉽게 전달할 수 있도록 지원한다.

메신저 모듈은 메티스와 질의자가 실시간으로 채팅을 통하여 질의를 주고 받기 위한 메신저 기능을 제공한다. 알포도 시스템은 platform-free한 메신저 개발과 확장성을 고려하여 메신저의 국제 표준 규격인 XMPP (extensible messaging and presence protocol)를 채용하여 개발되었다. XMPP는 2004년 IETF에 의해 표준안으로서 인정된 메신저 프로토콜로 이 규격을 따르는 메신저는 클라이언트 프로그램의 종류에 구애받지 않고 동작한다.

VoIP 모듈과 협업 모듈은 사용자와 메티스와의 커뮤니케이션의 효율을 향상시키기 위한 도구로 음성 통신을 이용한 지혜 교류와 화이트보드 등을 이용한 지혜 교류 기능을 제공한다.

6절 웹 인터페이스

알포도 시스템과 질의자와의 인터페이스를 위해서는 RIA(rich internet application) 플랫폼을 이용한 유연성 있는 사용자 인터페이스가 사용되었다. RIA 환경을 구축하기 위한 툴로, 개발 경험이 있는 Web Trustform과 Web 2.0 시대에 맞추어 새로 개발된 adobe사의 Flex 2.0 기술이 사용되었다.

웹 인터페이스에는 사용자-메티스 상호평가 모듈이 포함되어 있다. 상호평가 모듈은 질의자가 메티스의 성실도, 답변의 정확성 및 신속도 등 평가하는 기능을 제공한다. 사용자의 메티스 평가는 질의응답 세션이 종료된 후, 수행되며 사용자의 평가 결과를 메티스의 프로파일에 반영하여 양질의 메티스가 검색되도록 활용한다.

7절 구현 및 평가

알포도 시스템은 한신대학교 자연어처리연구실과 (주)H&B의 산학협력을 통하여 개발되었으며, 현재 베타

서비스를 <http://rpodo.com/beta.jsp>를 통해서 실시 중이다.

알포도 시스템의 성능 평가는 여러 가지 관점에서 실시 될 수 있다. 사용자의 질의어와 연관된 메티스 검색의 정확도와 재현율이 첫 번째 평가 기준이다. 질의어의 자동 분류 정확도, 메티스의 자동 갱신 모듈의 정확도와 이를 활용한 메티스 검색 성능의 향상 정도도 중요한 평가 기준이다. 하지만 위에서 설명한 모든 기준은 대량의 메티스 확보와 테스트베드 구축을 통하여 가능할 것이며, 이를 위해서는 다소간의 시간이 소요된다. 따라서 본 논문은 테스트베드를 활용한 객관적인 정량적인 평가는 향후 과제로 남겨두고 주관적인 평가를 수행하였다. 주관적인 평가로 측정된 것은 사용자의 질의에 적합한 메티스의 탐색 여부가 MOS(mean opinion score)값에 의해서 측정되었다. 측정 방법은 1,000명의 메티스를 대상으로 1,000건의 질의를 이용하여 수행되었다. 질의어를 이용하여 메티스를 검색하고, 검색된 메티스들에게 현재의 질문이 본인에게 전달된 것에 따른 의견(opinion)을 5점 척도로 평가하게 하였다. 측정 결과 평가 평균 3.92점의 만족도를 보이고 있음을 알 수 있었다.

8절 결론

웹 2.0 시대의 도래로 인터넷 사용자는 더 이상 정보의 사용자만이 아닌 정보의 생산자의 역할을 수행하는 프로슈머(prosumer)의 역할을 수행하게 되었다. 블로그 문화와 UCC의 활황으로 전문가 못지않은 정보와 지혜를 보유하고 생산하는 사용자가 급속히 증가하고 있는 추세이다. 제안하는 알포도 시스템은 급증하는 정보의 프로슈머에게 지혜를 생산, 공유, 유통 배포할 수 있는 새로운 인터넷 세상에서의 지혜 교류 시스템으로 활용될 것으로 기대한다.

알포도 시스템의 메티스는 그들이 가진 지혜를 나누는 댓가로 사이버 머니와 포인트 획득뿐만 아니라 명성(reputation)의 획득을 통하여 그들의 지혜 공유에 대한 댓가로 보상받게 된다. 명성이 있는 메티스는 그들이 활동하고 있는 블로그의 방문자를 급증시키는데 훌륭한 역할을 할 것이며, 블로그내의 메티스의 UCC가 유료로 판매되어 UCC 유통 시장의 활성화도 기여할 것으로 기대된다.

예부터 노익장이라는 말이 있다. 실버계층은 삶의 오랜 경험을 통해 얻은 지혜를 가진 우리 사회의 뿌리라 할 수 있다. 알포도 시스템은 실버 계층의 지혜를 사회

에 전달하고 교류하여 지혜의 사회 환원이라는 중요한 역할을 수행하며, 실버 계층은 지혜의 환원 과정을 통해서 사회에 공헌하는 자신을 통해서 자존감 회복과 사회와 교통할 수 있을 것으로 기대된다.

본 논문이 제안한 알포도와 유사한 인력기반의 질의응답시스템으로는 국내의 지식맨(<http://cs.jisikman.com>) 서비스, 국외의 차차(<http://chacha.com>), 비트와인(<http://bitwine.com>) 등을 들 수 있다. 지식맨 서비스는 상호성과 질의응답커뮤니케이션에 있어서 알포도와 다른 서비스이다. 차차 서비스와 비트와인은 미국에서 베타 오픈되어 서비스되고 있는 기술로 실시간성과 상호성에 있어서 알포도 시스템과 매우 유사하나 질의 자동 분류, 메티스의 지능적 갱신 및 자동 분류 기술에 있어서 알포도 시스템과 차별성이 있다고 할 수 있다.

단순 키워드 검색, 문장 검색, 의미 검색, 그리고 질의응답시스템의 과정을 통해서 발전해온 정보 검색 기술은 컴퓨터 알고리즘만을 이용하여 지능적 검색을 시도하였다. 지능적인 자동 질의응답 시스템의 연구는 지속되어야 하고 언젠가는 인간에 못지않은 지적인 자동 질의응답시스템의 개발이 기대된다, 하지만 인간과 컴퓨터 알고리즘이 공조하여 지능적인 질의응답이 가능한 알포도 시스템은 많은 시간과 노력을 동원하여 개발된 완전 자동 질의응답시스템보다 훨씬 지능적이고 효율적이며 자연스러운 질의응답 서비스를 제공할 수 있으리라 생각한다.

참고문헌

- [1] Baeza-Yates and Ribeiro-Neto, Modern Information Retrieval, Addison-Wesley, 1999.
- [2] Salton, G. and McGill, M.J., Introduction to Modern Information Retrieval, McGraw-Hill, New York, 1983.
- [3] <http://cs.jisikman.com>
- [4] <http://www.chacha.com>
- [5] <http://www.bitwine.com>
- [6] <http://www.naver.com>
- [7] <http://www.rpodo.com>