

퍼지 기반 컴포넌트 검색 인터페이스 구현

김선아 · 한군희 · 한정수

천안대학교 정보통신학부

Fuzzy based Component Retrieval Interface Implementation

Seon-Ah Kim · Kun-Hee Han · Jung-Soo Han

Division of Information and Communication, Cheonan University

요 약

본 논문은 컴포넌트의 재사용을 위한 퍼지 기반 컴포넌트 검색 인터페이스를 구현하였다. 컴포넌트 검색을 위해서는 클래스의 상속관계를 이용한 시소러스로 구축하였고, 이를 통하여 질의를 이용한 컴포넌트 검색이 가능하도록 하였다. 또한 검색결과는 우선순위로 보여줌으로서 질의에 대한 보다 빠른 검색이 되도록 하였다. 검색된 컴포넌트는 원시코드, 컴포넌트 정보, 클래스 다이어그램 등을 제공함으로써 효율적인 컴포넌트 재사용이 가능하도록 하였다.

1. 서론

컴포넌트 재사용에서 컴포넌트의 검색 방법은 매우 중요하다. 검색에 사용되는 기법에는 연관(Associative)검색, 확률(Probabilistic)검색, 매칭(Matching)검색, 클러스터링(Clustering)에 의한 검색, 불리언 논리(Boolean Logic)검색, 그리고 온라인 검색(Online Retrieval)등이 존재한다.[1] 본 논문에서는 객체지향 컴포넌트의 효율적인 검색을 위하여 클래스의 상속관계를 이용하여 개념을 분류하였고, 퍼지 논리를 적용한 객체지향 시소러스 검색시스템을 구축하였다. 시소러스 구축은 클래스의 상속관계를 이용하여 개념들 사이의 관계를 의미를 갖는 구조로 확장하였다.

따라서 본 연구는 기존의 정보 검색 시스템에서 사용한 시소러스 방법과는 다른 객체지향 컴포넌트를 잘 표현할 수 있는 시소러스를 구축하는데 그 목적을 두었다. 본 시스템은 검색 노이즈의 감소를

위해서 질의 확장의 임계치를 조절하여 효율성을 시뮬레이션 함으로써 최적의 검색 효율을 나타내는 확장 임계치를 설정하였고, 컴포넌트 검색을 위한 인터페이스와 질의 처리기를 구현하였다. 또한 사용자를 위하여 시소러스 정보를 제공하고 검색된 컴포넌트의 재사용을 위하여 소스 코드와 클래스 다이어그램을 제공할 수 있도록 하였다. 그리고 검색된 컴포넌트는 유사도 순으로 검색되기 때문에 원하는 컴포넌트를 효율적으로 찾을 수 있는 재사용성이 높은 컴포넌트 검색시스템이다.

2. 관련 연구

CRCS(Computing Reviews Classification Structure)[2]는 5단계 계층구조의 트리 형태를 가지며 1000여 개의 키워드로 구성되어 있으며 키워드 사이는 'IS-A' 관계를 사용하고 있다. 그러나 도

메인에 대한 지식이 없거나 정확한 키워드를 모를 경우에 이용자가 기술하는 개념을 보다 잘 적용시킬 수 있는 그룹화 방법이 필요하다. FIRMS (Fuzzy Information Retrieval and Management System)[3]은 퍼지 객체를 처리하기 위해서 개발된 시스템으로 퍼지 속성을 가진 객체의 표현방법을 제안하였다. FIRMS는 객체의 속성값을 퍼지 정도로 표현하였고, 퍼지 속성 값을 가지는 퍼지 도메인을 정의하였다. 그러나 유사 관계를 연결하기 위하여 용어들의 개념적인 정의를 해야 하는 어려움이 있다. 계층적 시소러스 시스템[4]은 계층적 분류를 위한 범주를 설정하고, 컴포넌트가 행위적 특성에 따라 분류되는 방법을 제안하였다. 컴포넌트 특성은 소프트웨어 디스크립터(software descriptor: SD)로부터 추출된다. 그러나 컴포넌트의 검색이 아닌 멤버함수와 파라미터를 이용한 검색이기 때문에 클래스가 증가할수록 멤버함수가 기하급수적으로 증가하여 노이즈가 많아진다는 단점이 있다.

3. 인터페이스 구조

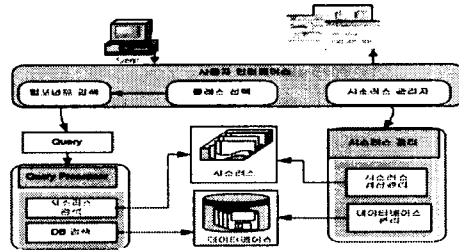
3.1 연구 방법

본 연구에서 컴포넌트 검색을 위한 시소러스 구축과 연구 방법은 클래스 간의 상속과 관련성에 따라 클래스 범주를 설정하고 클래스의 상속관계에 의한 개념 분류와 통계적 방법을 이용한 퍼지 시소러스를 구축하였다. 그리고 시소러스를 통한 질의 확장을 위한 최적의 임계치를 설정하였고, 질의 형성을 위해 클래스 시소러스 정보가 제공되며, 검색된 후보 컴포넌트에 대한 소스 코드와 클래스 다이어그램이 제공됨으로써 효율적인 컴포넌트의 재사용이 가능하도록 하였다.

3.2 검색 인터페이스 구조

(그림 1)은 검색 인터페이스의 구조를 보여준다. 사용자 인터페이스는 사용자로부터의 모든 입력과 출력을 처리하며, 클래스 선택, 컴포넌트 검색, 그리고 시소러스 관리로 구성된다. 컴포넌트 검색은 다시 질의 입력, 컴포넌트 검색 결과 리스트 출력, 컴포넌트 정보 부분으로 세분화되며, 각 컴포넌트에 대한 자세한 정보를 얻을 수 있도록 클래스 관계, 클래스 다이어그램, 그리고 소스코드를 지원한다. 또한 관리자는 데이터베이스와 시소러스의 갱

신을 이용하여 시소러스를 관리할 수 있도록 하였다. 데이터베이스는 유어의 테이블을 이용한 시소러스로 구축하였다.

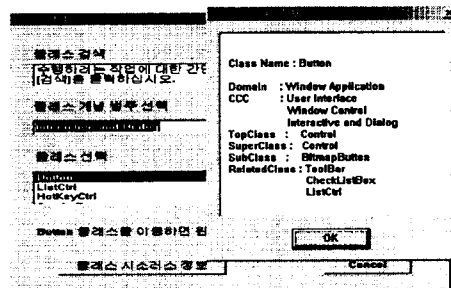


(그림 1) 검색 인터페이스 구조

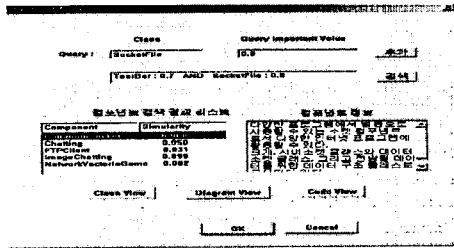
4. 검색 인터페이스 구현

4.1. 클래스 선택

본 시스템은 컴포넌트 검색에 대한 사용자의 질의 형성을 도와주기 위해 (그림 2)와 같은 클래스 선택 윈도우를 제공한다. 클래스에 대한 전문적인 지식이 풍부하지 않은 사용자를 위해 자연어 형식의 검색과 기능이나 특징에 해당하는 클래스 개념 범주를 선택함으로써 이에 해당하는 클래스를 제공한다. 제공된 클래스들을 질의로 선택할 때 각 클래스에 대한 텍스트 정보와 시소러스 정보를 참고함으로써 보다 효과적인 질의 형성이 가능하도록 하였다. (그림 2)에서 클래스 범주로 "Interactive and Dialog"를 선택하였을 때 이에 해당하는 클래스 리스트를 보여주고, 이 중 "Button" 클래스를 선택했을 경우에 "Button" 클래스의 정보를 보여준다. '클래스 시소러스 정보'는 각 클래스에 대한 시소러스 정보를 나타내는 윈도우로써 도메인, 클래스 상속관계와 유어 테이블에서의 유의값이 가장 높은 3개의 클래스에 대한 정보를 제공한다.



(그림 2) 클래스 선택



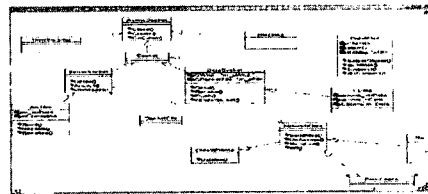
(그림 3) 컴포넌트 검색

4.2. 컴포넌트 검색

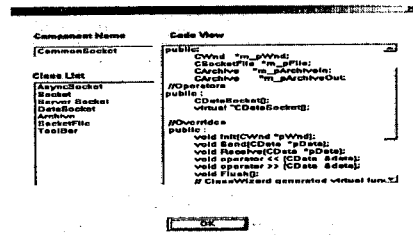
컴포넌트 검색은 질의 입력, 검색 결과 리스트, 컴포넌트 정보로 세분화된다. 질의 입력은 클래스 선택 윈도우로부터 선택된 클래스와 질의어 중요도(query important value)로 구성되며, 1개 이상의 질의어를 선택하길 원할 경우에는 '추가' 버튼을 눌러 다시 클래스 선택 윈도우를 구동시킬 수 있도록 하였다. 본 연구에서는 2개 이상의 질의에 대한 연결은 "AND"로 처리하였다. (그림 3)은 질의로 "ToolBar"와 "SocketFile"을 선택하고 각각의 질의 중요도를 0.7, 0.9로 주었을 때의 검색 결과를 보여 준다. 검색 결과는 시소러스에 의해 검색된 후보 컴포넌트들이 유사도 순으로 나타나 있고, 선택한 컴포넌트에 대한 자세한 정보가 컴포넌트 정보 부분에 제공된다. 또한 각 컴포넌트를 재사용하는데 있어서 필요한 정보를 얻을 수 있도록 클래스 관계 윈도우, 클래스 다이어그램 윈도우, 그리고 소스코드 윈도우를 제공한다. 많은 검색된 컴포넌트 중에서 적합한 컴포넌트를 선정하는 작업은 검색된 컴포넌트를 사용자가 이해해야 하기 때문에 많은 시간과 노력이 요구된다.

따라서 효율적인 재사용을 위해서는 검색된 컴포넌트들 중에서 사용자 요구에 가장 적합한 컴포넌트를 선정할 수 있도록 추천하는 기능이 필요하다. 클래스 다이어그램은 선택한 컴포넌트를 구성하는 클래스의 상속 구조를 표현한 것이며, 클래스 인스턴스 레벨에서의 컴포넌트 구조와 클래스 속성, 메소드 정보를 UML 다이어그램으로 제공하는 것인데, 이는 (그림 4)에서처럼 객체지향 모델링 툴인 Rational Rose와 연동되어 사용할 수 있도록 하였다. 외부 프로그램을 이용하는 검색은 사용자가 검색된 컴포넌트뿐 아니라 서로 연관된 컴포넌트를

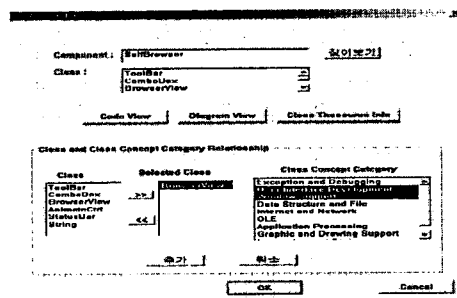
종합하여 참조할 수 있게 함으로써 검색된 컴포넌트를 쉽게 이해하여 선정하도록 도움을 줄 수 있다. (그림 5)는 검색된 컴포넌트의 관련 정보를 보여주고 있다. 이 정보에는 각 컴포넌트가 포함하고 있는 클래스들의 리스트와 코드를 보여주고 있다. 시소러스 관리 윈도우는 (그림 6)에서처럼 시소러스를 갱신하거나 데이터베이스에 접근하고자 할 때 사용된다. 시소러스는 인터페이스를 이용하여 자동으로 갱신되는데, 시소러스에 추가하고자 하는 클래스를 포함하는 컴포넌트를 선택한 후 추가하고자 하는 클래스와 클래스 개념 범주를 선택한다. 시스템은 클래스의 상속관계를 파악하여 추가될 클래스에 대해 클래스가 속한 개념 범주에 자동으로 할당해 준다. 이에 따라 클래스 정보, 소스 코드와 다이어그램 정보를 포함한 컴포넌트 정보가 데이터베이스에 저장된다.



(그림 4) 클래스 다이어그램



(그림 5) 컴포넌트 정보



(그림 6) 시소러스 관리

5. 결론

본 연구는 컴포넌트의 재사용을 위한 컴포넌트 검색 인터페이스를 구현하였다. 컴포넌트 검색을 위한 기술은 시소러스를 이용하여 구축하였고, 노이즈의 감소를 위해 질의 확장을 통한 컴포넌트 검색 인터페이스와 질의 처리기를 구현하였다. 또한 시소러스 정보를 제공하고 검색된 컴포넌트의 재사용을 위하여 소스 코드와 클래스 다이어그램, 시소러스 관리를 지원할 수 있도록 하였다. 그리고 검색된 컴포넌트는 유사도 순으로 검색되기 때문에 원하는 컴포넌트를 효율적으로 찾을 수 있다.

참 고 문 헌

- [1] Frakes and B.William "Information Retrieval-Data Structure and Algorithms," Prentice-Hall, 1992.
- [2] ACM, "The Full Computing Reviews Classification System," ACM, New York, 1992.
- [3] P. Subtil, N. Mouaddib and O. Foucaut, "A Fuzzy Information Retrieval and Management System and Its Applications," The Proceeding of the ACM Symposium on Applied Computing, pp.537-541. Feb. 1996.
- [4] E. Damiani, M. G. Fugini and C. Bellettini, "Aware Approach to Faceted Classification of Object-Oriented Component," ACM Trans. on Software Eng. and Methodology, Vol.8, No.4, pp.425-472. Oct. 1999.