

## WWW기반의 C 원시 코드 분석 도구의 개발\*

이명재 · 박종민  
컴퓨터 · 정보통신공학부

### <요약>

급격하게 발전하고 있는 웹기반의 환경에 맞추어 소프트웨어 공학기술은 변화하고 있다. 본 논문에서는 소프트웨어공학의 역공학과 UML 기술을 이용하고, 웹 프로그래밍 기술 중에서 자바 서블릿 및 애플릿을 사용하여, 웹상에서 C 원시 코드 파일을 분석하는 C 원시 코드 분석 도구의 설계 및 구현에 대해 논한다. C 원시코드 분석 도구는 사용자가 분석하기를 원하는 C 원시 코드를 웹 브라우저를 통하여 입력받아서 사용자에게 구조도 (Structure Chart)형태로 보여준다. 이 도구는 크게 분석기와 생성기로 나뉜다. 분석기는 C 원시 코드 파일을 입력받아서 프로젝트 단위로 저장하고, 역공학 도구의 일종인 cflow 와 cxref를 사용하여 C 원시 코드를 분석하여 데이터베이스에 저장한다. 생성기는 데이터베이스에서 자바 서블릿을 이용하여 자료를 추출하고, 자바 애플릿으로 그 결과를 사용자의 웹 브라우저에 구조도 형식으로 보여준다.

## Development of a WWW-Based C Source Code Analysis Tool

Myeong-Jae Yi, Jong-Min Park  
School of Computer Engineering and Information Technology

### <Abstract>

Due to the rapid growth and popularity of a WWW, the existing software engineering technologies are required to adapt to a new environment. In this paper we discuss the

---

\* 이 논문은 2000년 울산대학교의 연구비에 의하여 연구되었음

design and implementation of a documentation tool extracting automatically the design information from C source code files using reverse engineering technologies and UML.

The tool's outputs include the information about relationships between files and functions, between functions and functions, and between functions and global variables. To implement this tool, we used cxref and cflow that are reverse engineering tools, and java applet and servlet.

## 1. 서론

급격하게 발전하고 있는 웹기반의 환경에 맞추어 소프트웨어 공학기술은 변화하고 있다. 웹 기반 프로그램은 누구나 시간과 장소에 제한 없이 접근하여 사용할 수 있고, 특별한 설치프로그램이 필요 없으며, 개발자에게는 소프트웨어 업그레이드의 편리함을 제공한다는 장점을 가지고 있다. 결국, 사용자에게 좀 더 질 좋은 서비스를 제공해 줄 수 있다.

따라서, 본 논문에서는 웹기반 프로그램의 여러 가지 장점과 기존 소프트웨어공학의 장점을 이용하여 웹기반의 C 원시 코드 문서화 도구(WEBDOC-C)를 설계 및 구현하였다.

소프트웨어 유지보수는 소프트웨어 공학에서 매우 큰 비중을 차지하고 있다. 왜냐하면 소프트웨어 유지보수 비용은 전체 소프트웨어 개발 비용에 쓰이는 모든 노력의 60퍼센트 이상을 차지하고 있고, 이 비용은 계속 증가하기 때문이다. 또한, 프로그램 이해 분야는 소프트웨어 유지보수의 가장 근본적인 분야이다.[1]

소프트웨어를 설계하는 도구들 중에서 UML(Unified Modeling Language)은 대표적인 모델링 언어이다. UML은 표기하려는 대상을 다이어그램을 사용하여 나타낸다. 본 논문에서는 UML의 여러 가지 다이어그램 중에서 유즈-케이스 다이어그램(Use Case Diagram)과 순차 다이어그램 (Sequence Diagram)을 사용하여 전체적인 시스템 구조를 설계하였다.

본 논문의 구성은 다음과 같다.

2장에서는 역공학과 UML에 대해 기술하고, 3장에서는 WEBDOC\_C의 설계 및 구현 과정을 논의하며, 4장에서는 WEBDOC\_C의 사용법을 설명한다. 그리고 5장에서는 결론 및 향후 연구 과제에 대하여 서술할 것이다.

## 2. 역공학과 UML

### 2.1 역공학을 이용한 기법

역공학은 소프트웨어의 설계와 분석정보를 찾기 위하여 소프트웨어를 분석하는 과정이다. 본 논문에서는 역공학 도구의 일종인 Unix의 C 프로그램 분석 도구 cflow와 cxref를 사용하여 C 원시코드를 분석하였다. [2]

Cflow는 C 원시코드 파일을 읽어와서 함수와 함수사이의 상관관계를 콘솔상에 표준출력으로 나타내는 도구이다. 사용방법은 일반 콘솔상에서 cflow와 분석하고자 하는 C 원시코드 파일의 경로를 넣어주면 된다.

결과는 명령라인 밑으로 출력된다. 결과에서 제일 왼쪽의 숫자로부터 참조되어 main 함수가 start 함수를 호출하고, 다시 start 함수는 mysql\_connect 함수를 호출하는 관계가 표시되어 있다. 탭의 개수로 호출깊이를 표시하며 중괄호 안에 그 함수가 어디서 호출되었는지 파일명과 라인수로 표시되어있다.

```

1  main {refc.c 318}
2  start {refc.c 29}
3  mysql_connect{}
4  exitrr {refc.c 285}
5  mysql_error{}
6  fprintf{}
7  exit{}
8  mysql_select_db{}
9  fopen{}
.
.
.

```

Cxref는 C 원시파일에서 파일과 함수의 정보 및 전역변수에 대한 정보를 추출해준다.

Cxref는 C 원시코드파일을 입력받아서 표준출력과 LaTeX, HTML, RTF 및 SGML로 출력할 수 있다.

```

FILE : 'stacktest.c'

DEFINES : 'MAX' = 512
Defined : stacktest.c:1

VARIABLE : stack [Global]
Defined : stacktest.c:7
Type: int stack[512]

FUNCTION : init_stack [Global]
Defined : stacktest.c:11
Type: void init_stack
Arguments: void
.
.
.

```

Cxref는 파일명과 선언명령어와 전역변수 및 함수에 대한 정보를 추출해준다.

기본적인 사용법은 cflow와 같다. 콘솔 상에서 cxref와 원하는 C 원시 코드 파일을 차례로 입력하면 된다. 위의 결과에서 제일 위쪽부분인 “FILE” 부분은 C 원시 코드 파일명이며, “DEFINES”는 C 프로그래밍 명령어 중에서 정의 부분을 명시해준다. 여기서는 ‘MAX’라는 선언을 해주고 값 512를 넣었다는 것을 나타낸다.

전역변수는 “VARIABLE”로 표시하며 stack이라는 전역 변수는 “stack.c”의 11번째 라인에서 선언되었고, “int”라는 형을 표시한다. 함수는 전역변수와 유사하다. 함수명과 선언된 곳, 형과 인수값을 나타내준다.

우리는 위의 두 도구를 사용하여 C 원시코드 파일로부터 값을 입력받아서 나온 결과를 다시 분석하여 데이터베이스에 넣기 위한, 간단한 C 프로그램을 만들었다.

## 2.2 UML을 이용한 설계

소프트웨어를 설계하는 도구들 중에서 UML(Unified Modeling Language)은 객체지향 분석과 설계를 위한 대표적인 모델링 언어이다. 객체기술에 관한 국제 표준화 기구인 OMG(Object Management Group)에서는 이미 UML을 국제표준으로 인정했다.

UML은 표기하려는 대상을 다이어그램(Diagram)을 사용하여 나타낸다.[3] 우리는 UML의 여러 가지 다이어그램 중에서 유즈-케이스 다이어그램과 순차 다이어그램을 사용하여 전체적인 시스템 구조를 도식화하였다.

유즈-케이스 다이어그램은 요구분석의 첫 단계에서 주로 사용자의 요구를 기술하는데 사용된다. 사용자 요구 다이어그램은 시스템의 정적인 구조를 보여준다.

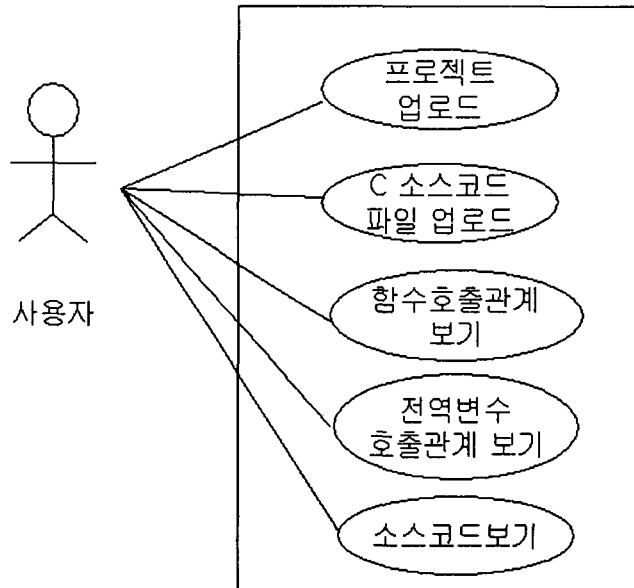


그림 1 시스템에 대한 유즈-케이스 다이어그램

위의 그림은 사용자가 할 수 있는 작업을 보여주고 있다. 사용자가 프로젝트를 업로드한 후 파일을 업로드하면 분석기는 자동으로 그것을 분석하여 데이터베이스에 저장하며, 생성기에 의해 만들어진 결과는 사용자의 요구 즉, 함수호출관계보기, 전역변수호출관계보기, 및 원시코드보기에 의해 구조도(Structure Chart)나 원시코드 형식으로 보여준다.

그림 2는 전체적인 흐름을 순차 다이어그램을 통하여 나타내고 있다.

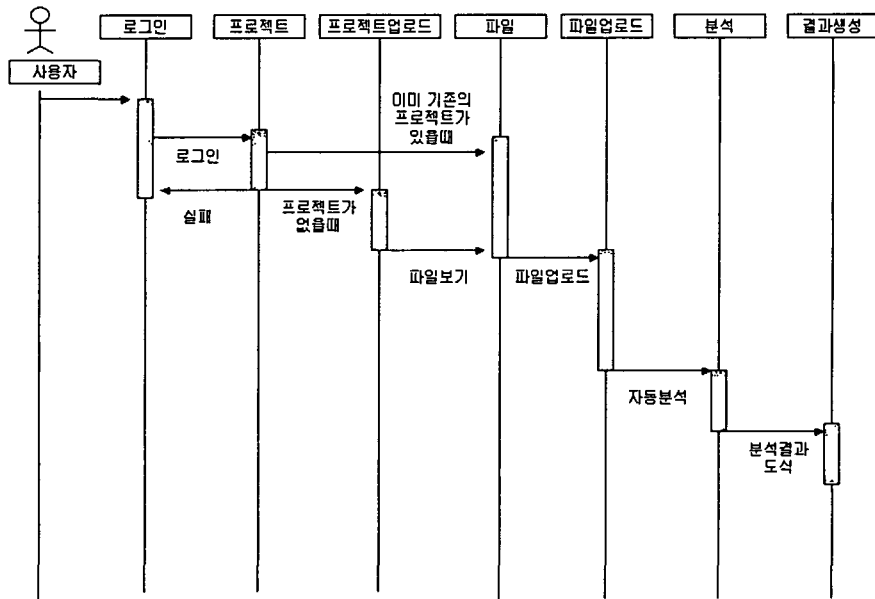


그림 2 순차 다이어그램을 이용한 설계화면

순차 다이어그램은 프로그램의 동적인 부분을 보여준다. 즉, 객체와 객체그룹사이, 객체와 객체사이, 객체그룹과 객체그룹 사이의 동적인 행위를 기술하게 된다.

특히 순차는 종좌표축으로 시간개념을 도입하고 횡좌표축으로 객체들을 나열하여 그 사이의 상호작용을 표시한다.[4]

사용자가 로그인을 하여 System에 성공적으로 접근하면 프로젝트 유무에 따라서 새 프로젝트를 업로드 할 것인지, 아니면 기존의 프로젝트를 사용할 것인지 결정하고, 파일을 업로드 하게된다. 업로드가 성공하면 분석기는 자동적으로 C원시코드를 분석하여 결과를 데이터베이스에 저장하고, 생성기는 사용자가 원하면 사용자에게 구조도 형식으로 보여지게 된다.

### 3. WEBDOC\_C 의 설계 및 구현과정

우리는 몇몇의 C 프로그램 및 자바 애플릿 및 서블릿과 PHP 스크립트를 사용하여 WEB 기반의 C 문서화 도구를 설계 및 구현하고 이름을 WEBDOC\_C라고 명명하였다.

WEBDOC\_C는 레드햇 리눅스 6.1 (커널 버전 2.3.6) 기반에서 아파치 웹서버 (Apache Web Server1.3.9)와 PHP 3.0으로 구성되고 서블릿 시스템으로 자카르타-톰캣(Jakarta-Tomcat)을 사용하였으며, 데이터베이스 시스템으로는 My-Sql을 사용하였다. [5]

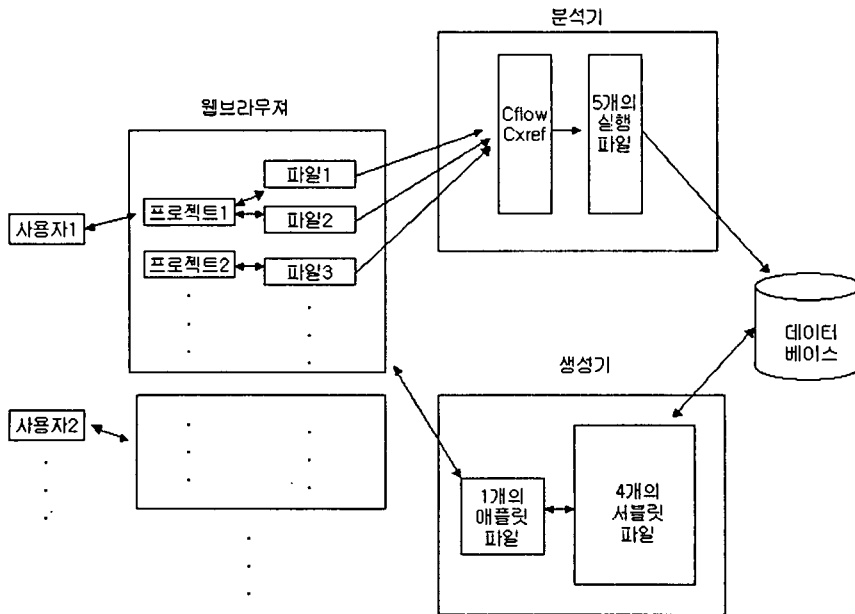


그림 3 전체 시스템 구조

WEBDOC\_C는 위 그림 3에서 보듯이 크게 두 부분으로 나뉜다.

사용자가 웹 브라우저로 WEBDOC\_C 시스템에 접근해서, 프로젝트를 생성하거나, 기존의 프로젝트를 선택하여, 원하는 C 원시코드파일을 업로드 하면, WEBDOC\_C는 C 원시코드파일을 분석하여 결과물들을 데이터베이스에 저장한다.

생성기는 데이터베이스로부터 분석된 결과물들을 가지고 자바 서블릿과 애플릿과의 통신을이용하여 웹 브라우저에 구조도 형태나, 원시코드 형태로 결과를 보여준다.

여기서 나오는 구조도는 함수와 함수사이의 호출관계 및 함수와 전역변수와의 관계를 보여주며, 원시코드는 함수나 전역변수가 어디에서 사용되고 있는가에 대한 정보를 원시코드상에서 보여주게 된다.

### 3.1 분석기

분석기는 역공학 도구인 Cflow 및 Cxref를 통하여 나온 결과를 가지고 WEBDOC\_C에 맞는 형식으로 바꿔서 데이터베이스에 저장하는 역할을 한다. 여기서 Cflow와 Cxref는 입력된 C 원시코드파일을 분석하여 알맞은 형식으로 변화시켜준다. 여기서 생성된 정보들을 바탕으로 5개의 간단한 C 프로그램을 이용하여 호출관계에 대한 정보를 데이터베이스에 저장한다. 데이터베이스는 8개의 테이블로 구성되어있다. 프로젝트와 사용자의 정보를 담고있는 3개의 테이블이 있으며, 2개의 테이블은 함수와 전역변수의 정보를 가지고 있고, 나머지 3개의 테이블은 함수와 함수사이의 호출관계와 함수와 전역변수와의 관계를 가지고 있다.

### 3.2 생성기

생성기는 데이터베이스로부터 자료를 가져와서 함수와 함수사이의 호출관계나, 전역변수와 함수와의 관계를 구조도 형식으로 나타내거나, 사용자가 업로드한 원시코드를 사용자의 웹브라우저에 보여주는 역할을 한다.

생성기는 4개의 자바 서블릿 클래스 파일과 1개의 자바 애플릿 클래스 파일로 되어 있는데 이것은 자바 애플릿의 보안모델이 "SandBox" 모델이기 때문이다. 이 모델은 애플릿은 잘 실행되지만, 클라이언트로의 자원접근은 강력하게 차단되어 있다. 따라서 자바 애플릿에서 직접 데이터베이스로 접근하는 것은 어렵기 때문에, 자바 서블릿을 사용하여 데이터베이스 접근하여 원하는 정보를 추출한다. 그 다음에 자바 서블릿과 애플릿간의 HTTP 기반 통신을 사용하여 4개의 자바 서블릿 클래스가 1개의 자바 애플릿 클래스로 정보를 전달하고, 자바 애플릿 클래스에서는 해시테이블과 3개의 구조체를 이용하여 정보를 가공하여 구조도나 원시코드 형태로 사용자의 웹브라우저에 보여준다.

## 4. WEBDOC\_C의 사용법

### 4.1 분석

사용자가 시스템에 접근하면 시스템은 먼저 사용자인증을 시작한다. 사용자가 적합하면 프로젝트 관련항목이 나타나는데, 만약 기존에 프로젝트 이름을 정하였으면 그것이 나타나고, 처음 사용자가 접근했거나, 프로젝트가 없을 경우는 새 프로젝트 업로드라는 항목이 나와서 프로젝트 업로드 할 수 있도록 해준다.

사용자가 프로젝트를 성공적으로 업로드 하거나, 기존의 프로젝트를 선택하였을 때, C원시코드파일에 대한 항목이 나온다. 만약 사용자가 이미 C원시코드파일을 업로드 하여서 분석을 했을 경우는 그것들에 대한 목록이 나타나며, 없을 경우 새로운 C원시코드파일을 업로드 하라는 항목이 나타난다. C원시코드 파일이나 C 헤더파일을 성공적으로 업로드 하여 자동분석을 시작하면, 3.1장에서 설명했듯이 분석기가 필요한 정보를 추출해서 데이터베이스에 저장시켜준다.

### 4.2 생성

그림 4는 WEBDOC\_C의 전체적인 사용 흐름을 나타내는 그림이다. 4.1의 분석 작업에서 DB에 저장된 자료들은 사용자가 원할 때 DB에서 추출되어 웹브라우저에 나타나게 된다.

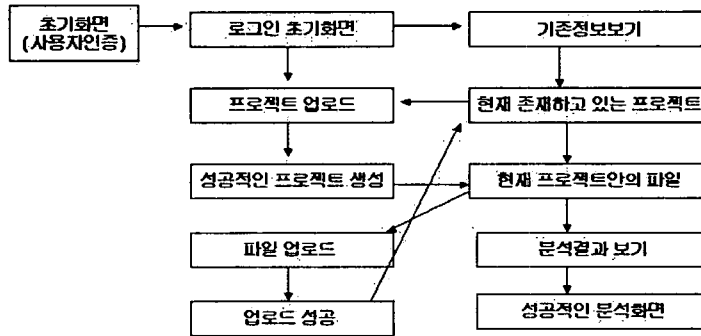


그림 4 WEBDOC\_C의 전체적인 사용 흐름

그림 5와 같이 생성기가 출력한 결과화면은 크게 2개의 프레임으로 나뉘진다. 왼쪽의 프레임은 명령어들이 위치하고 있으며, 오른쪽의 프레임은 결과를 구조도나 원시코드로 보여준다. 좌측상단의 리스트박스는 현재 프로젝트안의 모든 함수가 나타내며, 여기서 선택한 함수는 오른쪽 프레임의 결과창에 구조도로 나타나게된다.

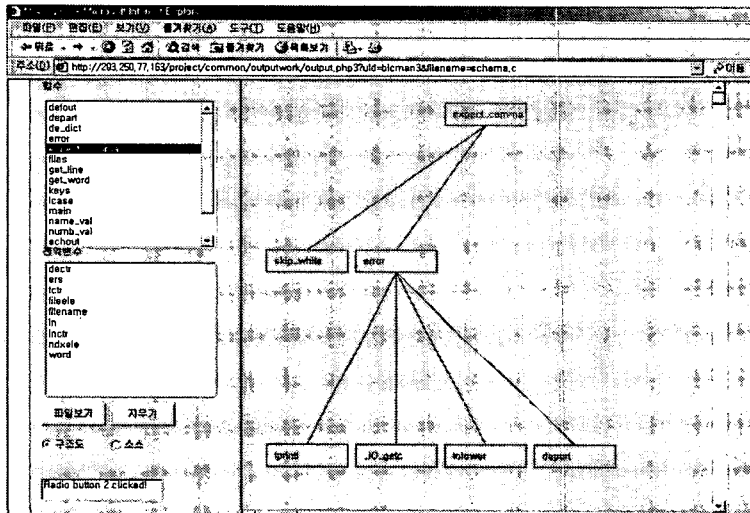


그림 5 함수 호출관계 출력

전역변수는 함수리스트 바로 밑에 나오는데, 함수리스트와 마찬가지로 전체 전역변수가 나타나며, 여기서 선택한 전역변수는 그림 6과 같이 오른쪽 결과창에 나타난다.



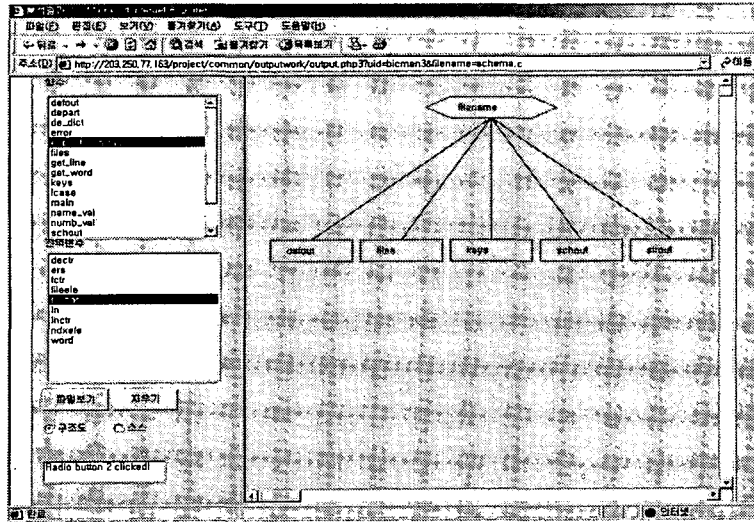


그림 6 전역변수 사용 관계 출력

전역변수 밑의 명령버튼들은 수행결과를 보는 버튼 하나와 오른쪽 결과창의 내용을 지워주는 역할을 하는 버튼 한 개로 이루어져 있다. 그 밑의 라디오버튼 2개중 하나는 구조도를 볼 때 선택하며, 또 다른 하나는 원시코드를 볼 때 선택한다. 원시코드를 선택하고, 변수나 함수중 하나를 선택하였을 때 오른쪽 결과창에는 그림 7과 같이 그 변수나 함수가 쓰인 라인이 붉은 색으로 표시되어 나타나게 된다.

그림7 원시 코드 보기 출력

## 5 결론 및 향후연구과제

본 논문에서는 역공학과 UML 및 웹 프로그래밍기법을 사용한 웹 기반의 C 원시 코드 문서화 도구의 설계 및 구현에 대해 논하였다.

이 도구의 개발을 위해 현재 전 세계적인 표준인 UML을 사용하여 시스템의 외부 행위와 전체적인 사용 흐름을 설계하였으며, 구현 방법으로는 다양한 웹 프로그래밍 기술을 사용하여 기존의 소프트웨어 공학 응용프로그램들과의 결합을 시도하였다. 앞으로 웹 기반의 기술이 지속적으로 발전할 것으로 예상되어, 웹 Metrics와 웹 유지보수(Web Maintenance) 기술을 포함한 웹 분석 (Web Analysis)을 연구할 계획이다.

## 참고문헌

1. Manna M., Maintenance Burden Beginning for a Remedy, Datamation, April 1993
2. AT&T Bell Laboratories, Unix System V Programmer's Manual,1985
3. Sinan Si Albir, UML in a nutshell, O'Reilly and Associates,1998
4. 송호중, 자바 개발자를 위한 UML contact J, 도서출판 대림, 2001
5. Randy Jay Yarger, George Reese & Tim King, My SQL & mSQL, O'Reilly and Associates, 2000
6. Chikofsky E.,Cross J., Reverse Engineering and Design Recovery: A Taxonomy, IEEE software, Vol 7. No.1, Jan. 1990
7. Ian Sommerville, Software Engineering, Fifth Edition, Addison-Wesley 1998.
8. SoftWare BlackSmiths Home Pages, CDOC Documentation Tool for C, C++ and Java, <http://www.swbs.com>
9. Tall Tree Software Company Home Pages, DocJet Source Code Documentation System, <http://www.tall-tree.com/>
10. Imagix Corporation Home Pages, Imagix 4D-reverse engineering and software documentation tool, <http://www.imagix.com/>
11. Tzerpos V.,Farmaner H.G, Web-Based Presentation of Hierarchic Software Architecture, Proc. Workshop on Software Engineering (on) the World Wide Web, Boston, May,19,1997