SHIVAJI UNIVERSITY

# Determinig Informational Navigational Transactional Intent of Web Queries

by

Author Name

A thesis submitted in partial fulfillment for the
Bachelor Of Engineering

in the
Department of Computer Science and Engineering
Sharad Institute of Technology College of Engineering,Yadrav-Ichalkaranji

April 2015

SHIVAJI UNIVERSITY

# *Abstract*

Department of Computer Science and Engineering
Sharad Institute of Technology College of Engineering,Yadrav-Ichalkaranji

Bachelor of Engineering

Determining the user intent of Web searches is a difficult problem due to the sparse data available concerning the searcher. By providing more accurate results depending on such query intents the performance of search engines can be greatly improved. We examine a method to determine the user intent underlying Web search engine queries. The classifications of informational, navigational, and transactional represent the type of content destination the searcher desired as expressed by their query. We implemented our classification algorithm and automatically classified a separate Web search engine transaction log of over a million queries submitted by several hundred thousand users. The user intent is generally vague or multi-faceted, pointing to the need to for probabilistic classification. We illustrate how knowledge of searcher intent might be used to enhance future Web search Engines.

# *Acknowledgements*

# Contents

# List of Figures

# List of Tables

# Chapter 1

# INTRODUCTION

## 1.1 Introduction of Project

Today there is vast growth in internet, to search a typical word on the internet we need search engines to search a information related to that topic and it gives or provide you all information related to that topic as well as other also. It is well known that with the increase of the information available on the Web, it is increasingly difficult for search engines to satisfy the user information need. Identification of web query intents has recently been becoming a hot point of the web information retrieval research area. Web search queries are the starting point to access the contents in the WWW for most of the users.[1] Capturing the user intent behind a query statement is crucial for any search engine and is equivalent to figuring out the category to which the query belongs to. So, we have to take look on improving the quality of a search engineâĂŹs results. User context elements like interests, preferences and intents are the main sources exploited in information retrieval approaches to fulfill user information needs. Using the user intent to improve the query specific retrieval search relies on classifying web queries into three types

1. Informational
2. Transactional
3. Navigational

The informational queries are queries that users want to learn the relevant information about a specific topic. The transactional queries are those that users supposed to carry out a transaction or interaction with the website returned by the search engines such as downloading music, playing online game, conducting online shopping, etc. The navigational queries refer to the queries that the users want to reach specific websites or webpages. Users normally do not like to remember a string of URL or often forget it so that they utilize search engines to help them to find these results.

# Chapter 2

# LITERATURE REVIEW

## 2.1 The Automatic Identification of a User Goal For a Web Query

S. Lovelyn Rose and K. R. Chandran [2] researchers had proposed the automatic identification of a user goal for a Web query. They have proposed two categories of effective features in identifying the goal of a query based on past user-click behavior and anchor-link distribution. They employed for the classification of the queries which return search results with yahoo directory search. The query to be classified is passed through the Yahoo directory search. The returned categories are referred to as the intermediate categories. The position and the frequency of occurrence of the intermediate categories are noted for a maximum of 50 search results. The intermediate categories with words matching with the target category are mapped using direct mapping. The remaining words are mapped using the path length. The target categories were finally ranked based on position, frequency and a combination of position and frequency. Limitation of this study is that experiment was conducted on a potentially-based dataset: queries from the CS department may show a technical bias and potentially work related. Some of the characteristics that they observed may not be true of user queries in general.

## 2.2 Classification of User Intent for Web Searching

Bernard J. Jansen [3] has defined and present a comprehensive classification of user intent for Web searching. The classification consists of three hierarchical levels of informational, navigational, and transactional intent. After deriving attributes of each, we then developed a software application that automatically classified queries using a

Web search engine log of over a million and a half queries submitted by several hundred thousand users. Our findings show that and compared the results from this manual classification to the results determined by the automated method. This comparison showed that the automatic classification has an more than 80% of Web queries are informational in nature, with about 10% each being navigational and transactional. In order to validate the accuracy of our algorithm, we manually coded 400 queries accuracy of 74%. Of the remaining 25% of the queries, the user intent is vague or multi-faceted, pointing to the need for probabilistic classification. We discuss how search engines can use knowledge of user intent to provide more targeted and relevant results in Web searching.

# Chapter 3

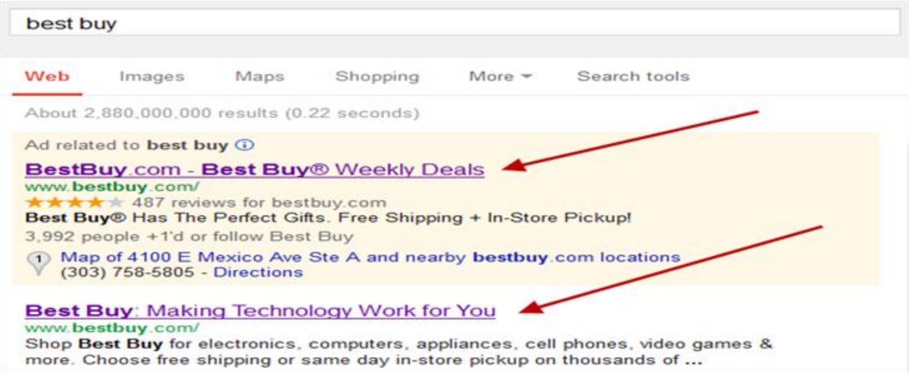
# TYPES OF SEARCH QUERIES

**Search Queries**
The words and phrases that people type into a search box in order to pull up a list of result's come in different flavors. It is commonly accepted that there are three different types of search queries:

1. Informational
2. Transactional
3. Navigational

## 3.1 Navigational Search Queries

A navigational query is a search query entered with the intent of finding a particular website or webpage. For example, a user might enter "youtube" into Google's search bar to find the YouTube site rather than entering the URL into a browser's navigation bar or using a bookmark. In fact, “facebook” and “youtube” are the top two searches on Google, and these are both navigational queries.



About 2,880,000,000 results (0.22 seconds)

Ad related to **best buy**

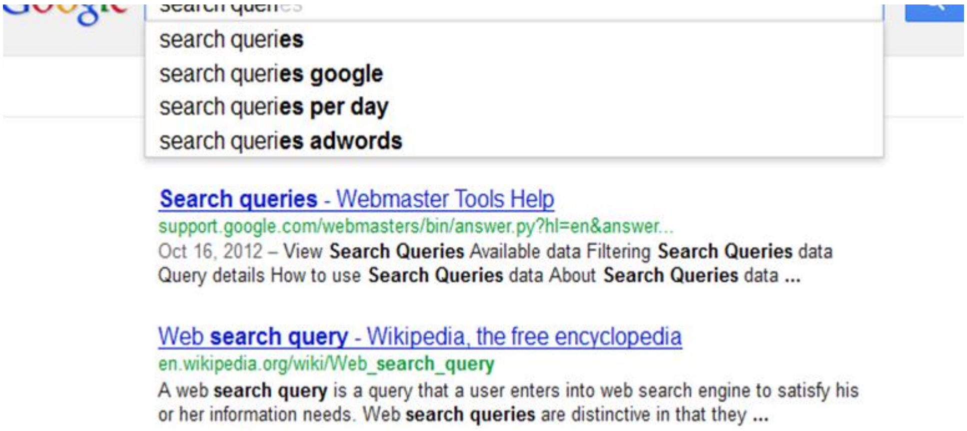
**BestBuy**.com - **Best Buy**® Weekly Deals
www.**bestbuy**.com/
★★★★★ 487 reviews for bestbuy.com
**Best Buy**® Has The Perfect Gifts. Free Shipping + In-Store Pickup!
3,992 people +1'd or follow Best Buy
Map of 4100 E Mexico Ave Ste A and nearby **bestbuy**.com locations
(303) 758-5805 - Directions

**Best Buy**: Making Technology Work for You
www.**bestbuy**.com/
Shop **Best Buy** for electronics, computers, appliances, cell phones, video games & more. Choose free shipping or same day in-store pickup on thousands of ...

FIGURE 3.1: Example of Transactional Search Queries

## 3.2 Informational Search Queries

Informational search queries as Queries that cover a broad topic for which there may be thousands of relevant results. When someone enters an informational search query into Google or another search engine, they are looking for information hence the name. They are probably not looking for a specific site, as in a navigational query, and they are not looking to make a commercial transaction. They just want to answer a question or learn how to do something.


search queries
search queries
search queries **google**
search queries **per day**
search queries **adwords**

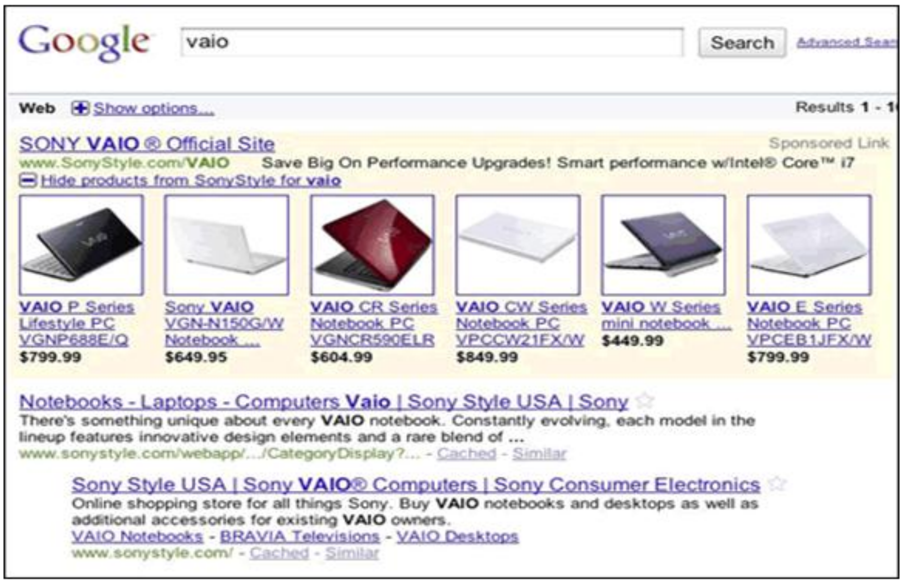
**Search queries** - Webmaster Tools Help
support.google.com/webmasters/bin/answer.py?hl=en&answer...
Oct 16, 2012 – View **Search Queries** Available data Filtering **Search Queries** data Query details How to use **Search Queries** data About **Search Queries** data ...

Web **search query** - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Web_**search_query**
A web **search query** is a query that a user enters into web search engine to satisfy his or her information needs. Web **search queries** are distinctive in that they ...

FIGURE 3.2: Example of Informational Search Queries

## 3.3 Transactional Search Queries

A transactional search query is a query that indicates an intent to complete a transaction, such as making a purchase. Transactional search queries may include exact brand and product names (like “samsung galaxy s3”) or be generic (like “iced coffee maker”) or actually include terms like buy, purchase, or order. In all of these examples, you can infer that the searcher is considering making a purchase in the near future, if they are not already pulling out their credit card. In other words, they are at the business end of the conversion funnel. Many local searches (such as “Denver wine shop”) are transactional as well.



SONY **VAIO** ® Official Site
Sponsored Link
www.SonyStyle.com/**VAIO** Save Big On Performance Upgrades! Smart performance w/Intel® Core™ i7
⊟ Hide products from SonyStyle for **vaio**

| | | | | | |
|---|---|---|---|---|---|
| **VAIO** P Series<br>Lifestyle PC<br>VGNP688E/Q<br>**$799.99** | Sony **VAIO**<br>VGN-N150G/W<br>Notebook ...<br>**$649.95** | **VAIO** CR Series<br>Notebook PC<br>VGNCR590ELR<br>**$604.99** | **VAIO** CW Series<br>Notebook PC<br>VPCCW21FX/W<br>**$849.99** | **VAIO** W Series<br>mini notebook ...<br>**$449.99** | **VAIO** E Series<br>Notebook PC<br>VPCEB1JFX/W<br>**$799.99** |

Notebooks - Laptops - Computers **Vaio** | Sony Style USA | Sony
There's something unique about every **VAIO** notebook. Constantly evolving, each model in the lineup features innovative design elements and a rare blend of ...
www.sonystyle.com/webapp/.../CategoryDisplay?... - Cached - Similar

Sony Style USA | Sony **VAIO**® Computers | Sony Consumer Electronics
Online shopping store for all things Sony. Buy **VAIO** notebooks and desktops as well as additional accessories for existing **VAIO** owners.
VAIO Notebooks - BRAVIA Televisions - VAIO Desktops
www.sonystyle.com/ - Cached - Similar

Figure 3.3: Example of Navigational Search Queries

# Chapter 4

# REQUIREMENT ANALYSIS

## 4.1 System Requirement

- Internet connection
- Browser
- Java1.6/1.7
- Eclipse 3.4/4.4 Standared
- Wordnet 2.1

## 4.2 Operating System Requirement

- Windows XP or later.

## 4.3 Tools and Technologies Requirement

### 4.3.1 Java 1.6/1.7

The Java Development Kit (JDK) is an implementation of either one of the Java SE, Java EE or Java ME platforms released by Oracle Corporation in the form of a binary product aimed at Java developers on Solaris, Linux, Mac OS X or Windows. The JDK includes a private JVM and a few other resources to finish the recipe to a Java Application.[2] Since the introduction of the Java platform, it has been by far the most widely used Software Development Kit (SDK).[citation needed] On 17 November 2006, Sun announced that it would be released under the GNU General Public License (GPL), thus making it free software. This happened in large part on 8 May 2007, when Sun contributed the source code to the OpenJDK.

The Java language has undergone several changes since JDK 1.0 as well as numerous additions of classes and packages to the standard library. Since J2SE 1.4, the evolution of the Java language has been governed by the Java Community Process (JCP), which uses Java Specification Requests (JSRs) to propose and specify additions and changes to the Java platform. The language is specified by the Java Language Specification (JLS); changes to the JLS are managed under JSR 901.

In addition to the language changes, much more dramatic changes have been made to the Java Class Library over the years, which has grown from a few hundred classes in JDK 1.0 to over three thousand in J2SE 5. Entire new APIs, such as Swing and Java2D, have been introduced, and many of the original JDK 1.0 classes and methods have been deprecated. Some programs allow conversion of Java programs from one version of the Java platform to an older one.After the Java 7 release, Oracle promised to go back to a 2 year release cycle. However, in 2013, Oracle announced that they would delay Java 8 by one year, in order to improve the Java security model.

### 4.3.2 Eclipse 3.4/4.4 Standared

This software is OSI Certified Open Source Software. Most of the Eclipse SDK is “pure” Java code and has no direct dependence on the underlying operating system. The chief dependence is therefore on the Java Platform itself. Portions are targeted to specific classes of operating environments, requiring their source code to only reference facilities available in particular class libraries (e.g. J2ME Foundation 1.1, J2SE 1.4, Java 5, etc). In general, the 3.6 release of the Eclipse Project is developed on a mix of Java 1.4, Java 5 and Java 6 VMs. As such, the Eclipse SDK as a whole is targeted at all modern, desktop Java VMs. Most functionality is available for 1.4 level development everywhere, and extended development capabilities are made available on the VMs that support them. contains a table that indicates the class library level required for each bundle.
Eclipse 3.6 is compatible with Eclipse 3.5 (and all earlier 3.x versions).
API Contract Compatibility: Eclipse SDK 3.6 is upwards contract-compatible with Eclipse SDK 3.5 except in those areas noted in the Eclipse 3.6 Plug-in Migration Guide . Programs that use affected APIs and extension points will need to be ported to Eclipse

SDK 3.6 APIs. Downward contract compatibility is not supported. There is no guarantee that compliance with Eclipse SDK 3.6 APIs would ensure compliance with Eclipse SDK 3.5 APIs. Refer to Evolving Java-based APIs for a discussion of the kinds of API changes that maintain contract compatibility.

Binary (plug-in) Compatibility: Eclipse SDK 3.6 is upwards binary-compatible with Eclipse SDK 3.5 except in those areas noted in the Eclipse 3.6 Plugin Migration Guide . Downward plug-in compatibility is not supported. Plug-ins for Eclipse SDK 3.6 will not be usable in Eclipse SDK 3.5. Refer to Evolving Java based APIs for a discussion of the kinds of API changes that maintain binary compatibility. Source Compatibility: Eclipse SDK 3.6 is upwards source compatible with Eclipse SDK 3.5 except in the areas noted in the Eclipse 3.6 Plug in Migration Guide . This means that source files written to use Eclipse SDK 3.5 APIs might successfully compile and run against Eclipse SDK 3.6 APIs, although this is not guaranteed. Downward source compatibility is not supported. If source files use new Eclipse SDK APIs, they will not be usable with an earlier version of the Eclipse SDK. Workspace Compatibility: Eclipse SDK 3.6 is upwards workspace compatible with earlier 3.x versions of the Eclipse SDK unless noted. This means that workspaces and projects created with Eclipse SDK 3.5 .. 3.0 can be successfully opened by Eclipse SDK 3.6 and upgraded to a 3.6 workspace. This includes both hidden metadata, which is localized to a particular workspace, as well as metadata files found within a workspace project (e.g., the .project file), which may propagate between workspaces via file copying or team repositories. Individual plug-ins developed for Eclipse SDK 3.6 should provide similar upwards compatibility for their hidden and visible workspace metadata created by earlier versions; 3.6 plug-in developers are responsible for ensuring that their plug-ins recognize metadata from earlier versions and process it appropriately. User interface session state may be discarded when a workspace is upgraded. Downward workspace compatibility is not supported. A workspace created (or opened) by a product based on Eclipse 3.6 will be unusable with a product based on an earlier version of Eclipse. Visible metadata files created (or overwritten) by Eclipse 3.6 will generally be unusable with earlier versions of Eclipse. Non compliant usage of API's: All non API methods and classes, and certainly everything in a package with “internal” in its name, are considered implementation details which may vary between operating environment and are subject to change without notice. Client plug ins that directly depend on anything other than what is specified in the Eclipse SDK API are inherently unsupportable and receive no guarantees about compatibility within a single release much less with earlier releases. Refer to How to Use the Eclipse API for information about how to write compliant plugins.

### 4.3.3 Wordnet 2.1

WordNet is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. WordNet is also freely and publicly available for download. WordNet's structure makes it a useful tool for computational linguistics and natural language processing. WordNet superficially resembles a thesaurus, in that it groups words together based on their meanings. However, there are some important distinctions. First, WordNet interlinks not just word forms"strings of letters"but specific senses of words. As a result, words that are found in close proximity to one another in the network are semantically disambiguated. Second, WordNet labels the semantic relations among words, whereas the groupings of words in a thesaurus does not follow any explicit pattern other than meaning similarity. The main relation among words in WordNet is synonymy, as between the words shut and close or car and automobile. Synonyms–words that denote the same concept and are interchangeable in many contexts–are grouped into unordered sets (synsets). Each of WordNet's 117 000 synsets is linked to other synsets by means of a small number of "conceptual relations". Additionally, a synset contains a brief definition ("gloss") and, in most cases, one or more short sentences illustrating the use of the synset members. Word forms with several distinct meanings are represented in as many distinct synsets. Thus, each form-meaning pair in WordNet is unique.

# Chapter 5

# SYSTEM DESIGN

## 5.1 Block Diagram

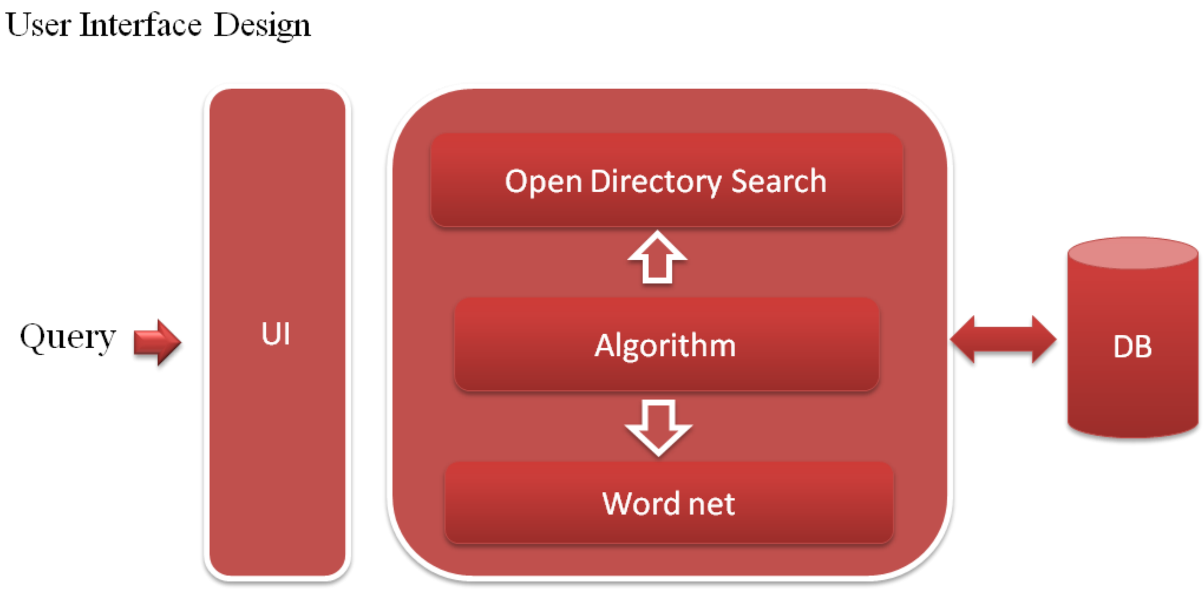


Figure 5.1: System Design

## 5.2 Algorithm

**Direct and Wordnet Based Mapping:**

1. Pass query through yahoo directory search
2. Retrieve a minimum of 50 search results if available
3. Record the position of the intermediate categories returned by yahoo directory search
4. Map the intermediate categories to the target category
    - Direct and Wordnet based Mapping
5. Rank the target categories
    - Rank

**Detaild Algorithm**
**Input:**
Intermediate categories ici retrieved by passing query through yahoo directory search.
Directcount initially 0, stores the number of times mapped directly.
First level target category initially 0 records the first level target category which is mapped.
Similarity measure target category count initially 0, stores the number of times mapped using wordnet.
**output:**
Required target category.

1. Step Tokenize iciâĂđs into icsij,s.

2. For every j in icsij, perform steps 3 to 10

3. Tokenize icsij,s into icwijk

4. For every k in icwijk which are not stopwords, perform steps 5 to 10

5. Increment Direct_count of the tcwijk which matches by Direct Matching DM(icwijk,tcwijk) for every i,j,k in tcwijk

6. For every k which has no result in DM(icwijk,tcwijk), perform steps 7 to 10

7. Calculate FLSM(icwijk,tcwijk). This would give an unambiguous viewpoint of the possible first category.

8. Increment First_level_target_category for tcwijk which returns Max(FLSM (icwijk,tcwijk))

9. Calculate SM(icwijk,tcwijk) between the intermediate category term and every tcwijk

10. Increment Similarity measure target category count of the tcwijk which returns Max(SM(icwijk,tcwijk))

11. Arrange tcsij in descending order of Direct_count, First level target category, Similarity measure target category count

12. If tcsij in the top position is of first level, then

    - Search for a high ranking subcategory of tcsij and thus decide the tci b. If no such tcsi is found, put the subcategory as "Others" Step 3: Else, take the corresponding first level category of tcsij

## 5.3 An Overview of UML

**The UML is a language for**

- Visualizing
- Specifying
- Constructing
- Documenting

THE UML LANGUAGE :

A language provides a vocabulary and the rules for combing words in that vocabulary for the purpose of the communication. A modeling language is a language whose vocabulary and rules focus on conceptual and physical representation of a system. A modeling language such as the UML is thus a standard language for software blueprints. In this context, specifying means building models that are precise, unambiguous, and complete. In particular, the UML addresses the specification of all the important analysis, design and implementation decision that must be made in developing and deploying a software intensive system. The UML is not a visual programming language, but its model can be directly connected to a verity of programming languages. This means that its possible to map from a model in the UML to a programming language such as java, cpp, or visual basic or even to tables in a relational database. Things that are best expressed graphically are done so graphically in the UML, whereas things that best expressed textually are done so in the programming language. A healthy software organization produces all sorts of artifacts in addition to raw executable code. These artifacts include requirements, architecture, design, source code, project plans, tests, prototypes, releases. The UML addresses the documentation of a systems architecture's and all of its details. The UML also provides for expressing requirements and for tests. Finally, The UML provides a language for modeling the activities of project planning and release management.

## 5.4 Goals of UML

The primary goals in the design of the UML were:

- Provide users with a ready-to-use, expressive visual modeling language so they can develop and exchange meaningful models. Provide extensibility and specialization mechanisms to extend the core concepts.
- Be independent of particular programming languages and development processes.Provide a formal basis for understanding the modeling language
- Encourage the growth of the OO tools market.
- Support higher-level development concepts such as collaborations, frameworks, patterns and components.
- Integrate best practices

## 5.5 A Conceptual model of UML

To understand the UML, you need to form a conceptual model of the language, and this requires learning three major elements: the UML's basic building blocks, the rules that dictate how those building blocks may be put together, and some mechanisms that apply throughout the UML. Once you have grasped these ideas, you will be able to read UML models and create some basic ones. As you gain more experience in applying the UML, you can build on this conceptual model, using more advanced features of the language.

### 5.5.1 Building Blocks of UML

The vocabulary of the UML encompasses three kinds of building blocks:

- Things
- Relationships
- Diagrams

These are the abstractions that are first-class citizens in a model; relationships tie these things together; diagrams groups interesting collections of things.

## 5.6 Diagrams in UML

A diagram is the graphical presentation of a set of elements, most often rendered as a connected graph of vertices (things) and arcs (relationships). You draw diagrams to visualizing a system from different perspectives, so a diagram is a projection into a system. For all but the most trivial systems, a diagram represents an elided view of the elements that make up a system. The same element may appear in all diagrams. In theory, a diagram may contain any combination of things and relationships. The views that comprise the architecture of software intensive system. For this reason, the UML includes following diagrams:

- Use case diagram
- Class diagram
- Sequence diagram
- Deployment diagram

## 5.7 Use Case Diagram

A use case diagram is a diagram that shows a set of use cases and actors and their relationships. A use case diagram is a just special kind of diagram and shares the same common properties as do all other diagram-a name and graphical contents.

### 5.7.1 Contents

Use case diagrams commonly contain

- **Use Case**

  Use case is a description of a set of sequence of actions that a system performs that yields an observable result of value to a particular actor. A use case is rendered as an ellipse with solid lines usually including its name.

- **Actors**

  An actor represents a role that an outsider takes on when interacting with the business system. For instance, an actor can be a customer, a business partner, a supplier, or another business system and every actor has a name.

- **Dependency, Generalization, and Association relationships.**

  A dependency is a semantic relationship between two things in which a change to one thing may affect the semantics of the other thing.

  A *generalization* is a relationship in which objects of specialized elements (the child) are substitutable for objects of the generalized element.

  An *association* is a structural relationship that describes a set of links, a link being connection among objects

  Like all other diagrams, use case diagram may contain notes and constraints.

### 5.7.2 Common Uses

Use case diagram typically contain in one of two ways.

- To model the context of the system

  Hear system involves drawing line around the whole system and actors outside of the system and interact with it.

- To model the requirement of a system

  Hear specifies what the system should do, independent of how that system should do.

### 5.7.3 Use Case Diagram

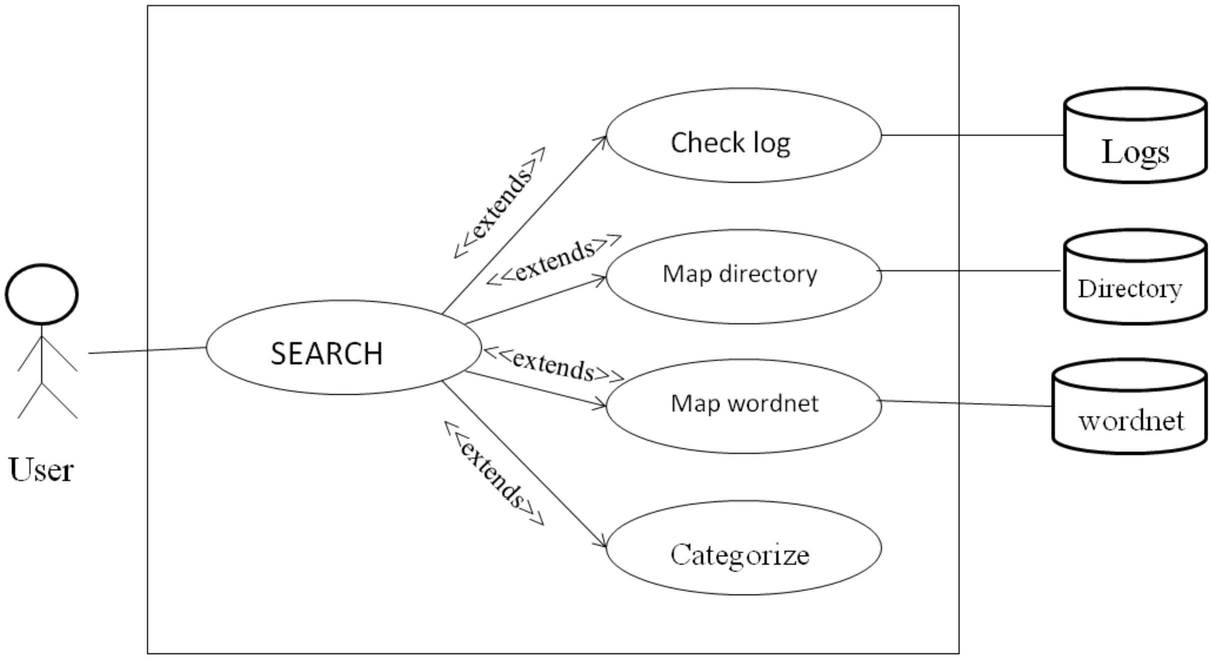


Figure 5.2: Use-Case Diagram

### 5.7.4 Use Case Scenarios

| **USE CASE** | **USE CASE SCENARIO** |
|---|---|
| Search Form | A USER/ACTOR need to insert query<br>Actors or USER |
| | **Alternate Flow of Event:**<br>User Insert Query in Search Form<br>User Click on Search button<br>If Query is not Found in Log Records<br>Then Query will go to Yahoo Directory Search<br>That query will mapped in Directory map<br>Directory Map will contain Categories<br>Categories are of three types<br>Navigational,Transactional and Informational |
| | **Alternate Flow of Event:**<br>User Insert Query in Search Form<br>User Click on Search button<br>If Query is not Found in Log Records<br>Then Query will go to Yahoo Directory Search<br>That query will mapped in Directory map<br>Directory Map will contain Categories<br>Categories are of three types |
| | **Alternate Flow of Event**<br>User Insert Query in Search Form<br>User Click on Search button.If Query is not Found in Log Records<br>Then Query will go to Yahoo Directory Search<br>That query will mapped in Directory map<br>Directory Map will contain Categories<br>Categories are of three types<br>If query is Navigational<br>It display the Website pages ,Social Sites etc. |
| | **Alternate Flow of Event**<br>User Insert Query in Search Form<br>User Click on Search button<br>If Query is not Found in Log Records<br>Then Query will go to Yahoo Directory Search<br>That query will mapped in Directory map<br>Directory Map will contain Categories |

Continued on next page

**Table 5.1 – continued from previous page**

| USE CASE | USE CASE SCENARIO |
|---|---|
| | Categories are of three types<br>If query is Transactional<br>It display to Purchase one of the Product etc. |
| | **Alternate Flow of Event**<br>User Insert Query in Search Form<br>User Click on Search button<br>If Query is not Found in Log Records<br>Then Query will go to Yahoo Directory Search<br>That query will mapped in Directory map<br>Directory Map will contain Categories<br>Categories are of three types<br>If query is Informational<br>It display information like Wikipedia etc. |

Table 5.1: Use Case Scenarios

### 5.7.5 Contents

**Sequence diagram commonly contains**

- Objects
- Links
- Messages

### 5.7.6 Definition and Overview

A *sequence* diagram is an interaction diagram that emphasizes the time ordering of messages. A sequence diagram shows a set of objects and the messages sent and received by those objects. The objects are typically named or anonymous instances of classes, but may also represent instances of other things, such as collaborations, components, and nodes. You use sequence diagrams to illustrate the dynamic view of a system. An Actor models a type of role played by an entity that interacts with the subject (e.g., by exchanging signals and data), but which is external to the subject (i.e., in the sense that an instance of an actor is not a part of the instance of its corresponding subject). Actors may represent roles played by human users, external hardware, or other subjects. Note that an actor does not necessarily represent a specific physical entity but merely a particular facet (i.e., "role") of some entity that is relevant to the specification of its associated use cases.

Sequence diagram have two features that distinguish them from collaboration diagrams.

- First, there is the object lifeline. An object lifeline is the vertical dashed line that represents the existence of an object over a period of time. So these objects are at the top of the diagram. With their lifelines drawn from the top of the diagram to the bottom
- Second, there is the focus of control. The focus of control is a tall, thin rectangle that shows the period of time during which an object is performing an action, either directly or through a subordinating procedure. The top of the rectangle is aligned with the start of the action; the bottom is aligned with its completion and also it can be marked by replay message.

### 5.7.7 Sequence Diagram

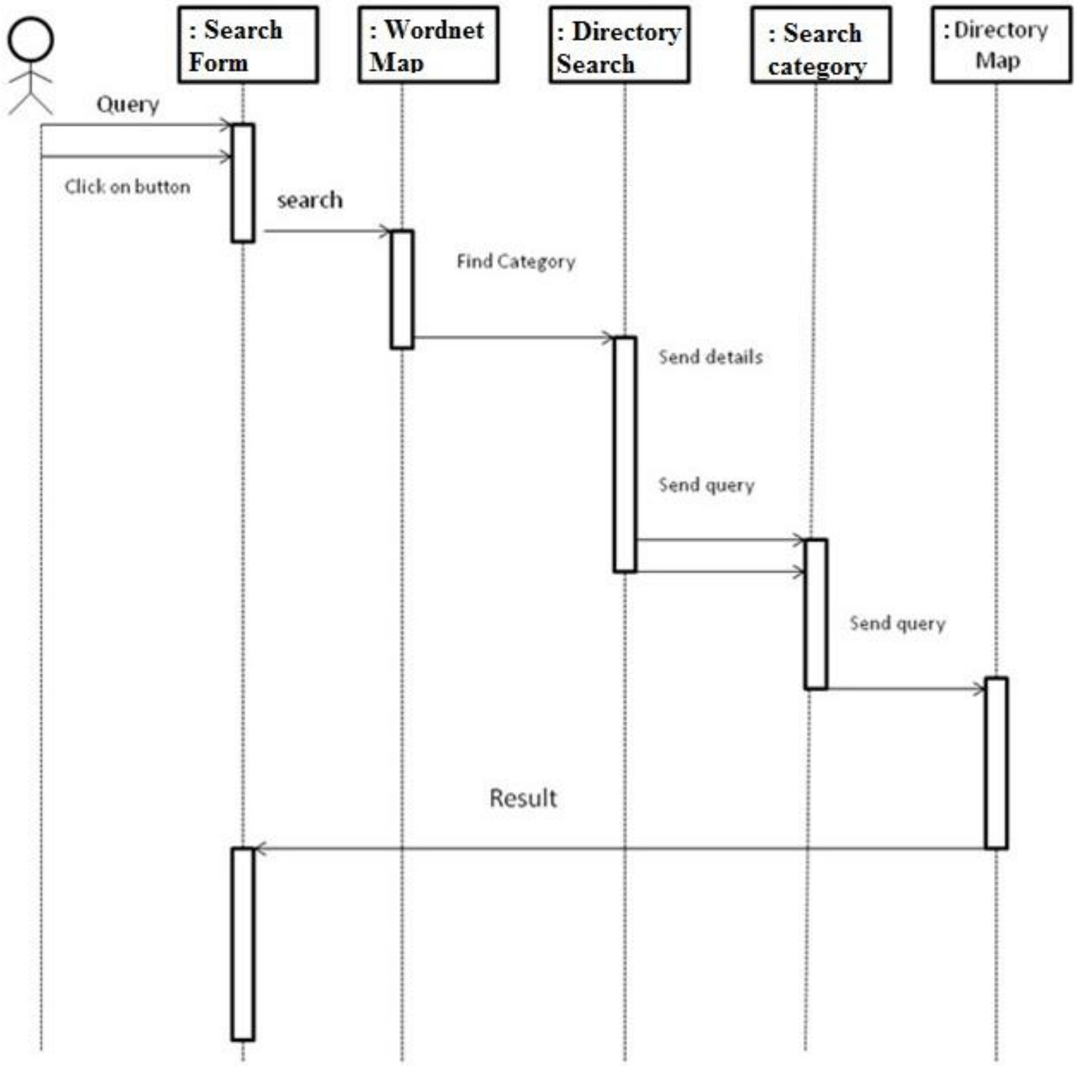


FIGURE 5.3: Sequence Diagram

## 5.8 Class Diagram

### 5.8.1 Contents:

Class diagram commonly contain the following things:

- Classes
- Interfaces
- Collaborations
- Dependency, generalization, and association relationships.

### 5.8.2 Definition and Common Uses

A class diagram is a diagram that shows a set of classes, interfaces and their relationships. Graphically, a class diagram is a collection of vertices and arcs. A class diagram will shares the same common properties as do all other diagrams. A class diagram is an illustration of the relationships and source code dependencies among classes in the Unified Modeling Language (UML). In this context, a class defines the methods and variables in an object, which is a specific entity in a program or the unit of code representing that entity. Class diagrams are useful in all forms of object-oriented programming (OOP). The concept is several years old but has been refined as OOP modeling paradigms have evolved. In a class diagram, the classes are arranged in groups that share common characteristics. A class diagram resembles a flowchart in which classes are portrayed as boxes, each box having three rectangles inside. The top rectangle contains the name of the class; the middle rectangle contains the attributes of the class; the lower rectangle contains the methods, also called operations, of the class. Lines, which may have arrows at one or both ends, connect the boxes. These lines define the relationships, also called associations, between the classes.

- Class: A definition of objects that share given structural or behavioral characteristics.
- Attribute: A typed value attached to each instance of a classifier.
- Operation: A method or function that can be performed by instances of a classifier

### 5.8.3 Class Diagram

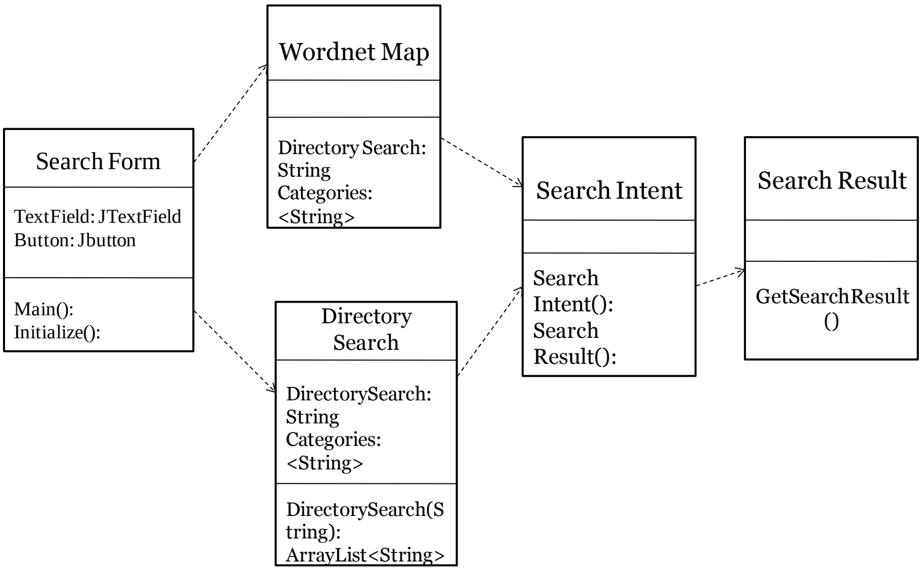


FIGURE 5.4: Class Diagram

## 5.9 Deployment Diagram

A deployment diagram shows the configuration of run time processing nodes and the components that live on them. Deployment diagram address the static deployment view of architecture. They are related to component diagrams in that a node typically encloses one or more components.

### 5.9.1 Nodes and Components

The UML provides a graphical representation of node. This canonical notation permits you to visualize a node apart from any specific hardware. Using stereotype this notation to represents specific kinds of processors and devices.

A *node* is a physical element that exists at run time and represents a computational resource, generally having at least some memory, and often processing capability. Graphically, a node is rendered as a cube. Every node must have a name that distinguishes it from other nodes. A name is a textual string. Components are things that participate in the execution of a system; nodes are things that execute components. Components represent the physical packaging of otherwise logical elements; nodes represent the physical development of components and components that things are executed by nodes. The

UML can often use stereotypes to specify new kinds of nodes that you can use to represent specific kinds of processors and devices. A *Processor* is a node that has processing capability, meaning that it can be executed by component. A*device* is a node that has no processing capability and general, represents something that interfaces to real world.

### 5.9.2 Deployment Diagram

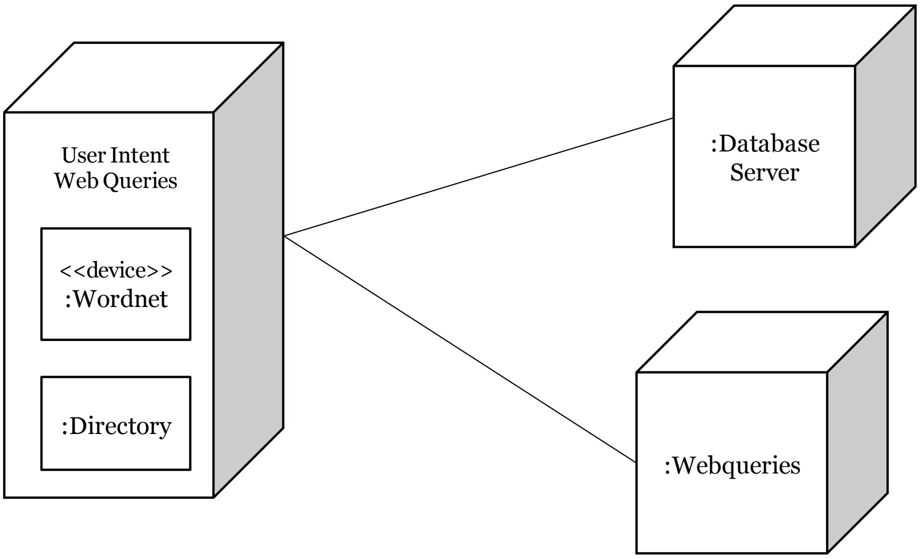


Figure 5.5: Deployment Diagram

# Chapter 6

# CODING

## 6.1 Installation of Tools

### 6.1.1 Java1.6/1.7

**Installation Steps:**

1. Open the Java SE Download page with this URL:[4] http://www.oracle.com/technetwork/java/java
2. Click the Download button under JDK in the "Java SE 7" section. You will see the "Java SE Development Kit 7 Downloads" page.
3. Click the download icon in " Windows x86 (32-bit) - jdk-7u3-windows-i586.exe" line.
4. Save jdk-7u3-windows-i586-p.exe to a temporary directory.
5. Double-click on jdk-7u3-windows-i586-p.exe to start the installation wizard.
6. The installation wizard will guide you to finish the installation.

### 6.1.2 Eclipse 3.4/4.4 Standared

**Eclipse 3.4 Installation :** Eclipse, both 32 and 64 bit can be downloaded from :[5] http://www.eclipse.org/downloads/

**Installation Steps:**

1. Go to http://www.eclipse.org/downloads/ to download a copy of eclipse. The first item listed is "Eclipse IDE for Java Developers (99 MB)." You want the 32-bit version, click on windows 32-bit.

2. Click on the green, downward pointing arrow. A download of eclipse-java-helios-win32.zip will start. Save the file to a convenient location, like your Desktop.

3. Find eclipse-java-helios-win32.zip from wherever you saved it. Then double-click it. It will open with the default windows extraction wizard or with an extraction tool like Winzip or Winrar. You will see a folder named "eclipse."

4. Drag the "eclipse" folder into your "Program Files." The easiest way to do so is to go to "Computer" (or "My Computer")–> "Local Disk (C:)" –> "Program Files" and drag the "eclipse" folder into the "Program Files" folder. Make sure that you do not drag it into a folder that's already within "Program Files." In other words, when you're done, the "Program Files" folder should have directly within it a folder named "eclipse." If Windows displays any dialog boxes asking for permission, click "allow" or "permit."

5. Double click the "eclipse" folder. You'll see an application named "Eclipse"; it has a purple icon with white horizontal stripes. Right-click it and press "Send To" -> "Desktop (Create Shortcut)." Now you will be able to launch Eclipse from your desktop.

6. Go ahead and open up Eclipse.

7. You'll see window, It will ask for the workspace .The workspace name will have your user name in place of "Zach." You can choose whatever place you want for your workspace, but it's easiest to just use the default you're given. I recommend that you just click the checkbox next to "Use this as the default and do not ask again" and then click "OK."

8. Welcome to the Eclipse IDE java for java developers window will appear. Click on the folded-over arrow on the right. You'll see the tooltip on the above screen when you move the mouse over the arrow. You won't see this screen again, even if you quit Eclipse and relaunch it.

9. Then you will see the window with Lots of buttons and panes.

10. You have now installed Eclipse.

## 6.2 Snippet

### 6.2.1 Classifier

This code will categories the the query into informational, navigational or navigational type.

**Transactional Quriese**

```
if(query.contains("movie")||query.contains("song")||query.
contains("lyric")||query.contains("download")||

query.contains("audio")||query.contains("video")||query
.contains("pictures")||query.contains("games")||

query.contains("movie")||query.contains("movie"))
                {
                        return true;
                }
                return false;
Informational Quriese

if(query.contains("what")||query.contains("how")||query.contains("who")|
("price")||query.contains("audio")||

query.contains("get")||(query.split(" ").length >3))
                {
                        return true;
                }
                return false;
```

### 6.2.2 Directory Search Map

This code extracts results from the yahoo directory.

```
ArrayList<String> Categories=null;
        String DirectorySearchURL =
   "http://dir.search.yahoo.com/search?fr=sfp&p=";
```

```
        String urls=null;
        public String getUrls()
        {
                return urls;
        }


                try {
                        yahoo = new URL(DirectorySearchURL +
   URLEncoder.encode(search));


                        HttpURLConnection uc;


                BufferedReader in;
                String inputLine;


                        uc =
   (HttpURLConnection)yahoo.openConnection();

                    uc.addRequestProperty("User-Agent",
   "Mozilla/5.0 (Windows NT 6.1; WOW64)
AppleWebKit/537.36 (KHTML, like

Gecko) Chrome/30.0.1599.101 Safari/537.36");

                    uc.addRequestProperty("Accept-Language",
   "en-US,en;q=0.8");


                     in = new BufferedReader(new
   InputStreamReader(uc.getInputStream(), "UTF-8"));


                                StringBuffer ba=new
   StringBuffer();
```

```
                                while ((inputLine = in.readLine())
!= null)
                                        {

ba.append(inputLine+"\n");
                                        }
                                        System.out.println("String
=="+ba.toString());
                                        StringBuilder urldata =
parserResults(ba.toString());
                                        setUrls(urldata.toString());
                                in.close();
                                } catch (IOException e) {
                                                // TODO
Auto-generated catch block
                                                e.printStackTrace();
                                        }
```

# Chapter 7

# TESTING

## 7.1 Testing

## 7.2 What is Software Testing

Software testing is the process of analyzing or operating software for the purpose of finding bugs. Testing can be described as a process used for reveling defects in software, and for establishing that the software has attained a specified degree of quality with respect to selected attribute.The fundamental objective of testing is to find defects ,as early as possible and get them fixed.

**Software Testing Process**

- Test Planning high level plans which list test objectives, test approach, measurement criteria along with test schedule and resources.
- Test Design create test cases, identify test cases for automation(if applicable),prioritize test cases and finalize test iterations.
- Test Implementation Create test scripts using automated testing tools.
- Test Execution Execute the test cases on the test environment and test reports.
- Test analysis Use test and project metrics to calculate key indicators. The data usually will be obtained from your defect tracking system.
- Postmortem reviews Discuss lessons learnt and identify strategies which will prevent such problems in future.

## 7.3 Test methods

### 7.3.1 Black Box Testing

It is also called as functional testing, it is the process of giving the input to the system and checking the output of the system.Without bothering about the system that how the system generates the output. It is also called as Behavior testing.

- Approach to testing where the program is considered as a Black Box.
- Testing based solely on analysis of requirements user specification, user documentation etc.
- The test cases are based on the specifications.
- Black box testing techniques apply to all levels of testing.
- Test planning and design can begin early in the software process.
- Tests are done from a users point of view.

### 7.3.2 White Box Testing

White box testing or structural testing considers facets like programming style, control method, source language, database design. A test for remote monitoring routine can be an example of structural test. This type of testing helps to uncover defects at structural level. The tests go below the top or functional layer to uncover the defects.

- Testing that takes into account internal structure and flow of a system or component.
- The testing is based on code structure or the algorithm.
- White box testing assumes that the procedural design and code is known to the tester.
- Obliviously test design can be done only after coding is complete.
- White box tests are inherently finite.

## 7.4 Test Cases and Test Data

- Test data are inputs that have been devised to test the system.
- Test cases are inputs and output specification plus a statement of the function under test.
- Test data can be generated automatically or real.

### 7.4.1 Test Cases

| TC ID | OBJECTIVE | PREREQUISITES | STEPS TO BE FOLLOWED | EXPECTED RESULT | ACTUAL RESULT | REMARK |
|---|---|---|---|---|---|---|
| 1 | Wordnet mapping | User must enter the query | Query passed to the wordnet | The wordnet will give correct meaning of query | Wordnet will give actual meaning | pass |
| 2 | Directory mappping | Pass Wordnet query result | Wordnet query will passed to yahoo directory | The yahoo directory will give similar results of entered query | The yahoo directory will give similar results of entered query | pass |
| 3 | Get Expected Results | Directory mapping | Check logs | Display results | Result be displayed | pass |
| 4 | Categories | Directory mapping | result is catagorised based on entered query | Categorised result has to be displayed | Categorised result displayed | pass |
| 5 | Iinformational Category | Directory mapping | result is catagorised based on entered query | Categorised result has to be displayed | informational Categorised result displayed | pass |
| 6 | Navigational Category | Directory mapping | result is catagorised based on entered query | Categorised result has to be displayed | Navigational Categorised result displayed | pass |

Continued on next page

**Table 7.1 – continued from previous page**

| TC ID | OBJECTIVE | PREREQUISITES | STEPS TO BE FOLLOWED | EXPECTED RESULT | ACTUAL RESULT | REMARK |
|---|---|---|---|---|---|---|
| 7 | Transactional Category | Directory mapping | result is catagorised based on entered query | Categorised result has to be displayed | Transactional Categorised result displayed | pass |

Table 7.1: Test Cases

# Chapter 8

# DEPLOYMENT

## 8.1 Deployment Steps

Deployment is next important procedure. As we discussed in this system require 3 Computers , Some tablet,1 wifi router for connectivity and printer.

- The computer is used as administrator which having all the authority. It must have the operating system and required softwares as mentioned in Software requirement.
- Connect the system to internet.
- Start Wordnet 2.1.It has to keep running while software is executing.
- Then run program by using eclipse.
- The search window will come.Enter your query or keyword int the text box thet you want to search.
- First It Extracts the information from the wordnet and deplay it.
- Then result is deplayed on the screen in the different categories.

## 8.2 Snapshots

**Search Form:-**

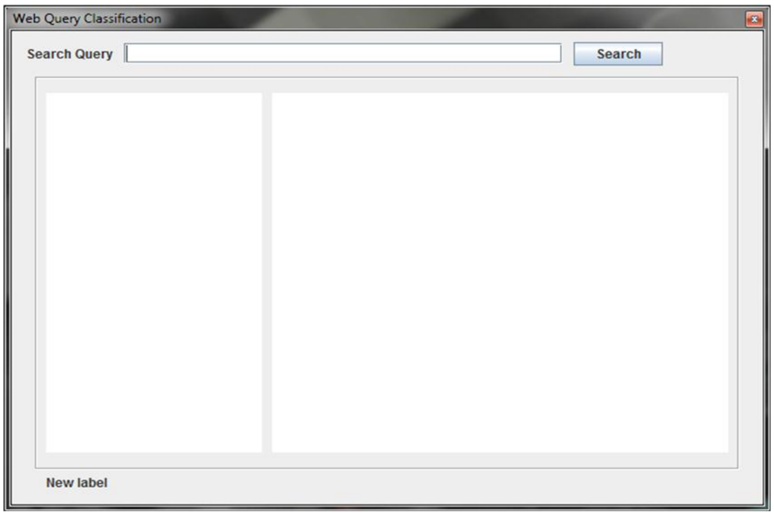
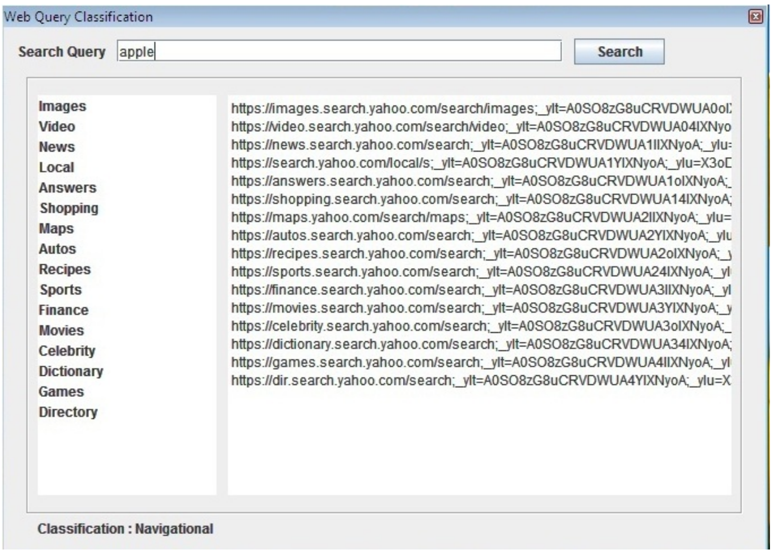


Figure 8.1: Search Form

When you run the project using Eclipse the search form will display. Put search query in the text box and click on Search button.

**Results on Search Form:-**



Search Query apple

```
Images
Video
News
Local
Answers
Shopping
Maps
Autos
Recipes
Sports
Finance
Movies
Celebrity
Dictionary
Games
Directory
```

```
https://images.search.yahoo.com/search/images;_ylt=A0SO8zG8uCRVDWUA0ol
https://video.search.yahoo.com/search/video;_ylt=A0SO8zG8uCRVDWUA04lXNyo
https://news.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA1lIXNyoA;_ylu=
https://search.yahoo.com/local/s;_ylt=A0SO8zG8uCRVDWUA1YlXNyoA;_ylu=X3oD
https://answers.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA1olXNyoA;_
https://shopping.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA14lXNyoA;
https://maps.yahoo.com/search/maps;_ylt=A0SO8zG8uCRVDWUA2lIXNyoA;_ylu=
https://autos.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA2YlXNyoA;_ylu
https://recipes.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA2olXNyoA;_y
https://sports.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA24lXNyoA;_yl
https://finance.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA3lIXNyoA;_yl
https://movies.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA3YlXNyoA;_y
https://celebrity.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA3olXNyoA;_
https://dictionary.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA34lXNyoA;
https://games.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA4lIXNyoA;_yl
https://dir.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA4YlXNyoA;_ylu=X
```

Classification : Navigational

Figure 8.2: Results on Search form

When you click on search button the information is extracted from the wordnet and desired results are displayed category wise.

**Web Page:-**

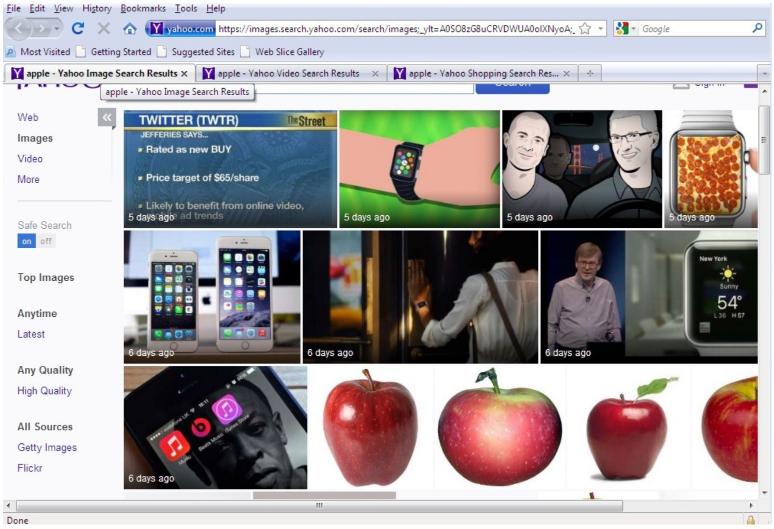
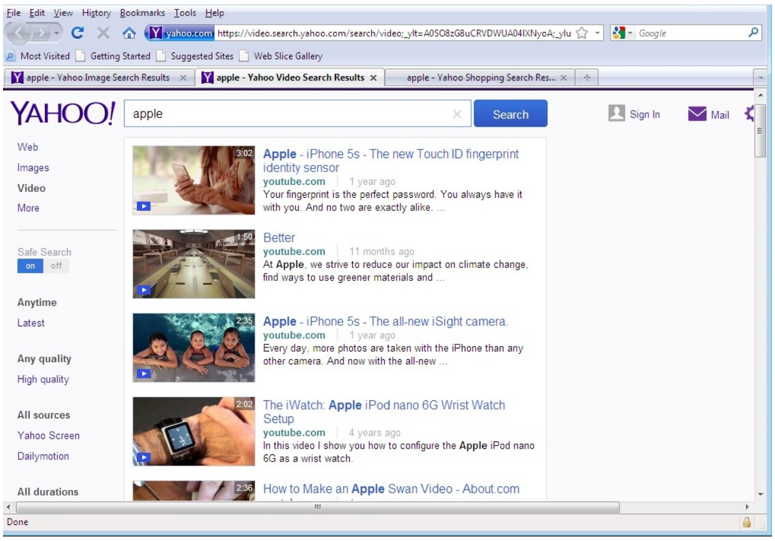


FIGURE 8.3: Result of Image Search

When you copy the URL of image and Enter into the Brower the related information will desplay.

**Web Page:-**


File Edit View History Bookmarks Tools Help
yahoo.com https://video.search.yahoo.com/search/video;_ylt=A0SO8zG8uCRVDWUA04IXNyoA;_ylu
Google
Most Visited
Getting Started
Suggested Sites
Web Slice Gallery
apple - Yahoo Image Search Results
apple - Yahoo Video Search Results
apple - Yahoo Shopping Search Res...
YAHOO!
apple
Search
Sign In
Mail
Web
Images
Video
More
Safe Search
on
off
Anytime
Latest
Any quality
High quality
All sources
Yahoo Screen
Dailymotion
All durations
3:02
1:50
2:35
2:02
2:36
Done

Apple - iPhone 5s - The new Touch ID fingerprint identity sensor
youtube.com | 1 year ago
Your fingerprint is the perfect password. You always have it with you. And no two are exactly alike. ...

Better
youtube.com | 11 months ago
At Apple, we strive to reduce our impact on climate change, find ways to use greener materials and ...

Apple - iPhone 5s - The all-new iSight camera.
youtube.com | 1 year ago
Every day, more photos are taken with the iPhone than any other camera. And now with the all-new ...

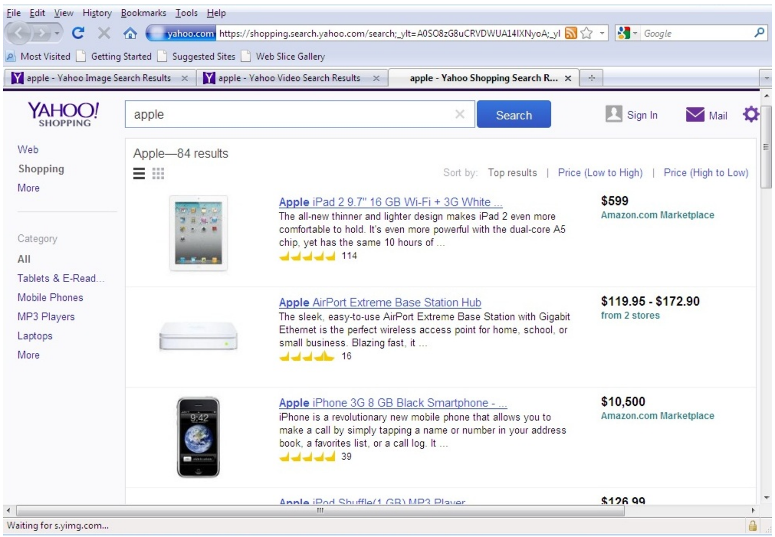
The iWatch: Apple iPod nano 6G Wrist Watch Setup
youtube.com | 4 years ago
In this video I show you how to configure the Apple iPod nano 6G as a wrist watch.

How to Make an Apple Swan Video - About.com

FIGURE 8.4: Result of video Search

When you copy the URL of video and Enter into the Brower the related information will desplay.

**Web Page:-**


File Edit View History Bookmarks Tools Help
yahoo.com
Google
Most Visited
Getting Started
Suggested Sites
Web Slice Gallery
apple - Yahoo Image Search Results
apple - Yahoo Video Search Results
apple - Yahoo Shopping Search R...
YAHOO! SHOPPING
Search
Sign In
Mail
Web
Shopping
More
Category
All
Tablets & E-Read...
Mobile Phones
MP3 Players
Laptops
More
Sort by: Top results | Price (Low to High) | Price (High to Low)

https://shopping.search.yahoo.com/search;_ylt=A0SO8zG8uCRVDWUA14IXNyoA;_yl

apple

Apple—84 results

Apple iPad 2 9.7" 16 GB Wi-Fi + 3G White ...
The all-new thinner and lighter design makes iPad 2 even more comfortable to hold. It's even more powerful with the dual-core A5 chip, yet has the same 10 hours of ...
114
$599
Amazon.com Marketplace

Apple AirPort Extreme Base Station Hub
The sleek, easy-to-use AirPort Extreme Base Station with Gigabit Ethernet is the perfect wireless access point for home, school, or small business. Blazing fast, it ...
16
$119.95 - $172.90
from 2 stores

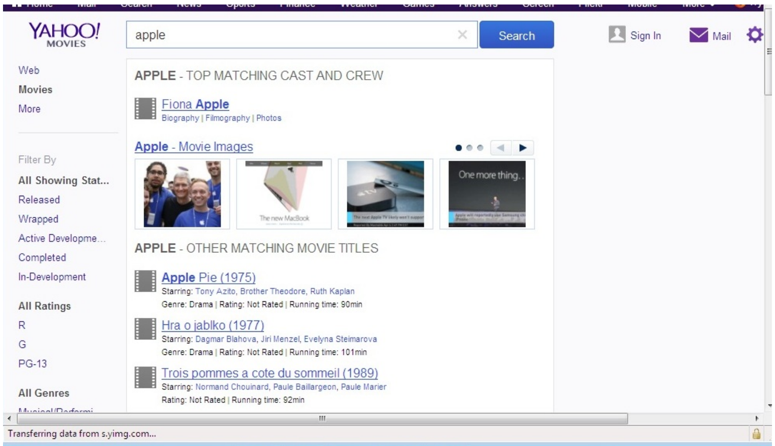
Apple iPhone 3G 8 GB Black Smartphone - ...
iPhone is a revolutionary new mobile phone that allows you to make a call by simply tapping a name or number in your address book, a favorites list, or a call log. It ...
39
$10,500
Amazon.com Marketplace

Waiting for s.yimg.com...

FIGURE 8.5: Result of Shopping Search

When you copy the URL of Shopping Search and Enter into the Brower the related information will desplay.

**Web Page:-**



apple

APPLE - TOP MATCHING CAST AND CREW

Fiona Apple
Biography | Filmography | Photos

Apple - Movie Images

APPLE - OTHER MATCHING MOVIE TITLES

Apple Pie (1975)
Starring: Tony Azito, Brother Theodore, Ruth Kaplan
Genre: Drama | Rating: Not Rated | Running time: 90min

Hra o jablko (1977)
Starring: Dagmar Blahova, Jiri Menzel, Evelyna Steimarova
Genre: Drama | Rating: Not Rated | Running time: 101min

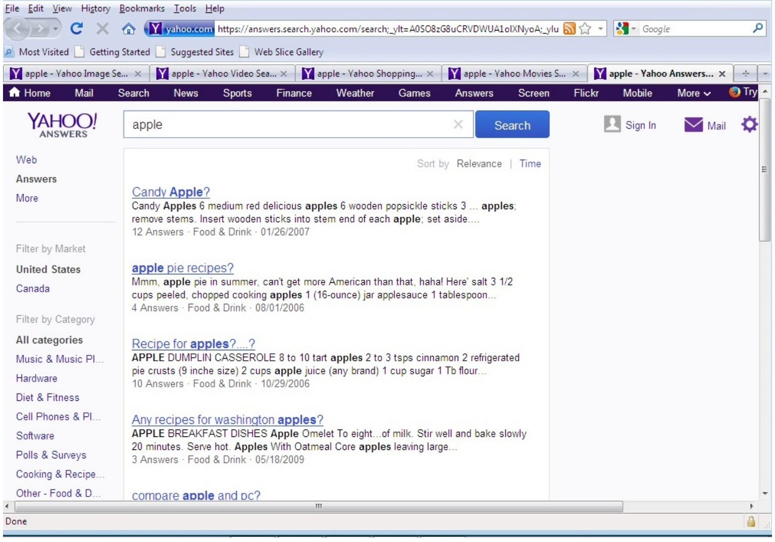
Trois pommes a cote du sommeil (1989)
Starring: Normand Chouinard, Paule Baillargeon, Paule Marier
Rating: Not Rated | Running time: 92min

Transferring data from s.yimg.com...

FIGURE 8.6: Result of Movie Search

When you copy the URL of movies and Enter into the Brower the related information will desplay.

**Web Page:-**



Candy **Apple**?
Candy **Apples** 6 medium red delicious **apples** 6 wooden popsickle sticks 3 ... **apples**; remove stems. Insert wooden sticks into stem end of each **apple**; set aside....
12 Answers · Food & Drink · 01/26/2007

**apple** pie recipes?
Mmm, **apple** pie in summer, can't get more American than that, haha! Here' salt 3 1/2 cups peeled, chopped cooking **apples** 1 (16-ounce) jar applesauce 1 tablespoon...
4 Answers · Food & Drink · 08/01/2006

Recipe for **apples**?....?
**APPLE** DUMPLIN CASSEROLE 8 to 10 tart **apples** 2 to 3 tsps cinnamon 2 refrigerated pie crusts (9 inche size) 2 cups **apple** juice (any brand) 1 cup sugar 1 Tb flour...
10 Answers · Food & Drink · 10/29/2006

Any recipes for washington **apples**?
**APPLE** BREAKFAST DISHES **Apple** Omelet To eight...of milk. Stir well and bake slowly 20 minutes. Serve hot. **Apples** With Oatmeal Core **apples** leaving large...
3 Answers · Food & Drink · 05/18/2009

compare **apple** and pc?

Figure 8.7: Result of Answers Search

When you copy the URL of Answers and Enter into the Brower the related information will desplay.

# Chapter 9

# CONCLUSION

## 9.1 Conclusion

We presented a method for classifying queries into informational, navigational and transactional queries. Identifying the category of the web search query is the major challenge in returning the right set of web search results and in inserting the most appropriate advertisement along with the web search results. We have demonstrated an automated method that can successfully classify web queries based on user intent. We have created framework for understanding the underlying goals of search, and have demonstrated that the framework can be used to associate goals with queries given limited information. We have highlighted characteristics of Web queries based on user intent. These characteristics were derived from an examination of web queries from multiple search engine transaction logs.

# Bibliography

[1] Dayong Wu, Yu Zhang, Shiqi Zhao, and Ting Liu. Identification of web query intent based on query text and web knowledge. In *Pervasive Computing Signal Processing and Applications (PCSPA), 2010 First International Conference on*, pages 128–131. IEEE, 2010.

[2] S Lovelyn Rose and KR Chandran. Web knowledge and wordnet based automatic web query classification. *International Journal of Computer Applications*, 17, 2011.

[3] Bernard J Jansen, Danielle L Booth, and Amanda Spink. Determining the informational, navigational, and transactional intent of web queries. *Information Processing & Management*, 44(3):1251–1266, 2008.

[4] Oracle. URL http://www.oracle.com/technetwork/java/javase/downloads.

[5] Eclipse 3.6. URL http://www.eclipse.org/downloads/.