A Web Site Navigation Engine

Mark Levene, Richard Wheeldon, Jon Bitmead
NavigationZone, Department of Computer
ScienceUniversity College London, U.K.

1   Introduction

Often users navigating (or "surfing") through a web site "get lost in hyperspace", when they lose the context in which they are browsing, and are unsure how to proceed in terms of satisfying their original goal. The unresolved problem in web site usability, of assisting users in finding their way, is termed the *navigation problem*. (See [LL00] for a survey and critique on the navigation problem.) This problem is becoming even more acute with the continuing growth of web sites in terms of their structure, which is becoming more complex, and the vast amount information they intend to deliver. In contrast users are not willing to invest time to learn this structure and expect the delivery of the relevant content without delay.

To tackle this problem we are developing a navigation system for semi-automating user navigation which builds *trails* of information, i.e. sequences of linked pages, which are relevant to the user query. The preferred trails are presented to the user in a tree-like structure which they can interact with. This is in sharp contrast to a search engine which merely outputs a list of pages which are relevant to the user query without addressing the problem of which trail the user should follow. We discuss the architecture of the navigation system and give a brief description of the navigation engine and user interface.

2   Architecture of the Navigation Engine

The architecture of the navigation system is shown in Figure 1. the *user interface* executes on top of a conventional browser such as Microsoft Internet Explorer or Netscape Navigator. It obtains the preferred trails for navigation, given a user query, from the *navigation engine* and requests pages for display from the Web site via a proxy.

The navigation engine consists of two main modules: (i) the information retrieval module, and (ii) the best trail module. The information retrieval module does conventional information retrieval over Web pages combined with a page ranking algorithm that takes into account the potential of a Web page as a starting point for navigation. The best trail module computes the preferred trails for navigation given the input query. The algorithm is adaptive in the sense that it dynamically searches for the preferred trails by mimicking a user navigation session and scoring the trails as they are expanded according to the topology of the Web site. (See [LL00] for some detail on the algorithmic process of used by the navigation engine to compute the preferred trails given a user query.)

The navigation engine interacts with a *Web case*, which is a database for storing the details of the pages of the Web site in preprocessed form. The *software robot* is an offline process which is responsible for the creation of the Web case.
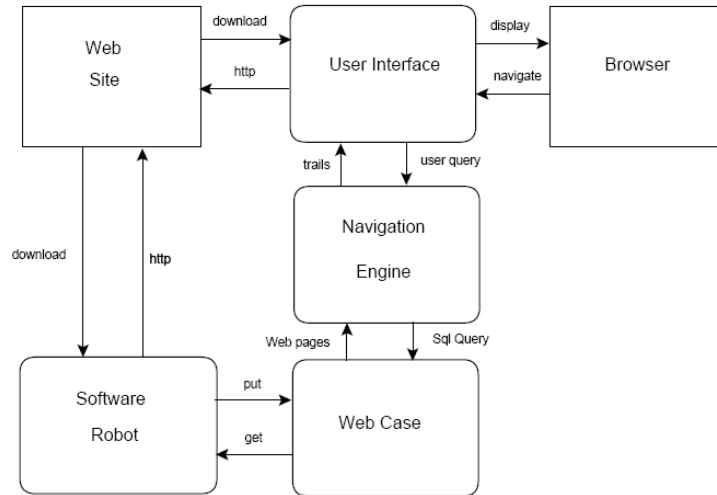
Figure 1: Navigation engine architecture

3   User Interface

The user interface includes two main mechanisms: (i) the *navigation tool bar* and (ii) the *navigation tree window*. The navigation tool bar comprises a sequence of URLs. The first two URLs displayed on the navigation tool bar are the last two URLs of the pages that the user browsed, thus providing a history mechanism. The next URL on the navigation tool bar is the current URL identifying the page the user is currently browsing. The URLs that follow the current URL on the navigation tool bar are the consecutive URLs of the best trail found from the current URL. All URLs are clickable and cause the navigation tool bar to be updated accordingly. The navigation tree window displays the preferred trails given the user query, organized in the form of a tree structure with the trails being ranked from the most preferred, according to their score. The user interacting with the navigation tree window can select any URL on one of the preferred trails causing it to be the current URL. The navigation tool bar, the navigation window and the browser window are all synchronized according to the current URL. The mechanisms of the user interface provide the user with guidance and context throughout a navigation session, given the input query. The user interface can be embodied in a Web site as a navigation mechanism complementing or replacing a Web site search engine.

References

[LL00]   M. Levene and G. Loizou. Web interaction and the navigation problem in hypertext. In A. Kent, J.G. Williams, and C.M. Hall, editors, *Encyclopedia of Microcomputers*. Marcel Dekker, New York, NY, 2000. To appear.