

A Browser for Browsing the Past Web

Adam Jatowt¹, Yukiko Kawai², Satoshi Nakamura³, Yutaka Kidawara³ and Katsumi Tanaka¹

¹Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku,
606-8501 Kyoto, Japan
{adam, tanaka}@dl.kuis.kyoto-u.ac.jp

²Kyoto Sangyo University
Motoyama, Kamigamo, Kita-ku,
603-8555 Kyoto, Japan
kawai@cc.kyoto-su.ac.jp

³National Institute of Information and Communications Technology
3-5 Hikaridai, Seikacho, Sorakugun,
619-0289 Kyoto, Japan
{gon, kidawara}@nict.go.jp

Abstract

We propose a framework for a past web browser – an interface for visualizing and navigating the Past Web. Past Web is considered as a set of past snapshots of web pages that can be downloaded from a series of web archives. The browser shows past page versions one by one along the time line in a passive way. Changes that occurred in histories of pages are animated by blinking and appearing/disappearing effects. The browser enables navigation in spatio-temporal structure of the past web. Additionally, query-based filtered browsing is enabled.

Main concepts

1. Past web vs. current web (live web)

- *Past web* is the current web as it looked like in the past

2. Meta-archive approach – simultaneous retrieval of past page versions from different past web repositories for improving the precision of reconstruction of page histories

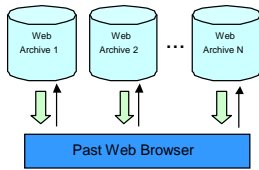


Fig2. Meta-archive data collection method

3. Two browsing types:

- *Horizontal browsing*: browsing the same page along the timeline
- *Vertical browsing*: browsing by following links on the page

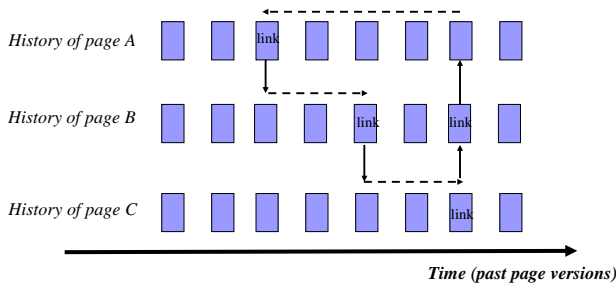


Fig3. Horizontal (dashed arrows) and vertical (solid arrows) type of browsing

4. Time-consistent and page-version-consistent back and forward buttons

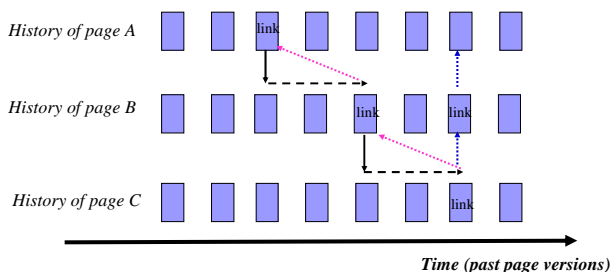


Fig4. Time-consistent (blue arrows) and page-version-consistent (pink arrows) navigation by pushing respective back buttons

5. Passive browsing – displaying past page versions in horizontal browsing one by one from certain starting time point like in a movie

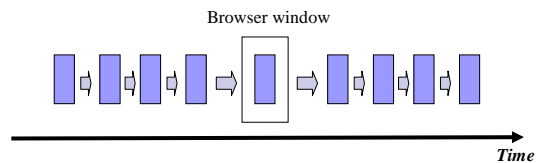


Fig5. Passive browsing concept

6. Change detection and animation – detecting changes between consecutive page versions during horizontal browsing and highlighting them by changed background colour and animation

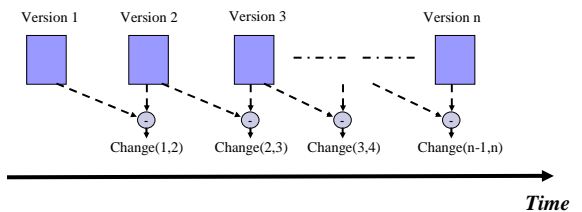


Fig6. Change detection in page history

7. Automatic jumps – automatic skipping of periods without any changes in horizontal browsing

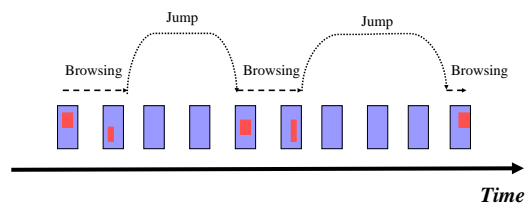


Fig7. Automatic jumping during horizontal browsing (changes shown as red rectangles)

8. Query-based filtered browsing – combination of search and browsing by displaying only changes containing user-specified query



Fig1. Browser interface