Improving Internet Interaction: from theory to practice

by

Russell Beale

Research Lecturer
School of Computer Science
University of Birmingham
Edgbaston
Birmingham
B15 2TT
email: r.beale@cs.bham.ac.uk
telephone: +44 (0) 121 414 3729
fax: +44 (0) 121 414 4281

Improving Internet Interaction: from theory to practice

The internet is a medium for education, entertainment, communication and personal expression. User behaviour has developed three main modalities for using this medium effectively - searching, browsing and monitoring, which are supported to different degrees by conventional tools. Understanding the nature of the interaction allows us to design and implement a system called Mitsukeru to support browsing behaviours, whilst retaining the freeform movements between other interaction styles. The system uses agent-based modelling and look-ahead to provide informative yet non-intrusive guidance to the user, and is described in detail.

Keywords: Internet, Agents, Interaction Modelling, User Modelling, Search, Browsing, Mitsukeru, Bayesian Statistics, DHTML

1 Understanding Users

Internet use is now common amongst a wide cross-section of society, as more and more people discover its abilities to inform, entertain, empower and aid communication. We have taken a principled approach to addressing the issue of improving user's interaction with the internet. Firstly, analysis of internet behaviour has been undertaken, leading to a categorisation of user behaviour. This basic description (not grand enough to qualify for a theory) can then form the basis for the design of enhanced systems to improve interaction.

Initial experiments with users involved talk-aloud evaluations, observation, and web log analysis, to build a picture of how people utilise the internet in their everyday activities. Users ranged from students and academics to family members, schoolchildren and office workers. People were monitored in how and why they used the internet over a period of between one and three weeks. This observation and reflection on user behaviour on the internet leads us to propose that there are three essential forms of behaviour whilst online; searching, browsing, and monitoring. These behaviours relate to how much the user desires the specific information, and whether they know where to find it or not. These terms can have a multitude of meanings, and so we define them more explicitly.

Searching is the process of trying to locate something specific. It relies on using search engines and other resources that take you to specific pages or areas, which either match your needs or not. The search terms used are refined and new searches done, or the current information is further filtered until the user finds what they are looking for, or not. Searching tends to be a focussed activity, occasionally interleaved with other tasks or internet activities, but usually strongly goal-directed, until the information required is located.

Browsing is the more random perusal of information presented, often accomplished by reading pages that are not directly relevant, or by following links or interesting threads of topic that reveal themselves. Browsing is often observed as users follow the links suggested by a site's navigational architecture. Browsing is not directed towards a specific goal. There will be short-term goals, usually relating to things of immediate interest or relevance, but these can easily change as other items are seen and investigated. It is certainly true that the diversity and easy accessibility of material on the internet means that users often drop into browsing behaviour when they undertaking a searching task - it is an easy distraction. Browsing can be seen as using the internet for entertainment purposes - a resource to be explored for its own sake, not to answer a pressing issue.

The final behaviour is that of *monitoring* - the repeated visitation of a web page. In these cases, the web pages are often dynamic, and monitoring behaviour can most usually be seen in people trying to obtain stock quotes, weather information, or news. They have a trusted page or site that they return to again and again to update themselves with the relevant information. This monitoring behaviour can occur over very different timescales (even for the same user) - for example, many people utilise the BBC website for weather information and visit it a couple of times a week; stock quotes for day traders are revisited every 15 minutes or so during active trades, then may be ignored for a month or more.

These three types of behaviour are very different, and are often combined in other tasks. For example, in deciding to buy a dishwasher, a user may well search for pages that rate different dishwashers. They read the reviews and decide which to get, but in the process they browse off and learn how a dishwasher works, and what the stock price for the two leading electrical retailers is. They then monitor an on-line store, waiting for the sale to start so that they can then buy their chosen item online at the cheapest possible price.

Deconstructing internet activity into these behaviours does allow us to better understand user behaviour, but it is important to realise that the intermixing of these styles is an inherent part of internet-based interaction.

2 Internet Complexities

To appreciate why using the internet is more complex than we might initially imagine, we need to understand more about the context and nature of the interaction. The complexity has to be more than simply the vastness of the internet resource - libraries can have huge quantities of information in them too. Some blame for the difficulty must lie with the unordered, unstructured nature of the internet - and yet with items often linked to related items, there should also be some benefit here, and in any case the massive computing resources most people can call upon ought to have some ability to impose some sort of order onto this chaos. There is more to it than that - part of the problem is in how people perceive, use and react to the internet itself.

In the non-electronic world, searching for information usually takes place in one of a few specific locations, most notably in a library (personal, community or commercial). Situated within that specific context, the task at hand is defined and restricted by the resources available. Organised index cards provide a global view of the catalogued and ordered information, allowing a search to have a defined starting point. If the aim is to find a specific atomic item - a book - then the index will allow the user to head straight there. If the aim is to find information on a specific subject, which may be contained within one book or more likely spread throughout many, then the search is developed through refinement and feedback until the correct stack and shelves and then books are found. The physical structure of the library also offers subtle supporting assistance to users. It supports physical memory - where things are and their spatial relationships - and visual memory - what a book looks like, how big or how old it is. This reality of the library - its architecture, layout and quiet air - all lend themselves to the focussing of the goal of retrieving knowledge and information. Precisely because of this, a library is not the chosen location to meet friends, or watch a movie, or catch a radio programme. Because the library is task-specific, the thought of these things is laughable.

However, in the electronic world, we use the internet for a whole host of things. We use it as our information resource, which we wish to search and find things in. We also use it as an entertainment resource, and a communication tool. It therefore doesn't "feel" much like a library. Because the internet has to support this multiplicity of roles, it is unlikely to be particularly good at one of them if that is at the expense of the others. So from a practical perspective, the fact that it impacts our lives in many more areas than a library does makes it harder for it to fulfil the information retrieval tasks so effectively. This generality of usage also means that, from a psychological perspective, the user is not working in such a situated, focussed context, and so their searching behaviour is often mixed in with diversions into other forms of behaviour [1]. This easy distraction also makes the effective finding of information difficult.

Coupled with this diversity of content is the new functionality of the internet; cross-linked information, multimedia, smaller atomic units, keyword and meaning-based search abilities, and so on. Because of these, users are modifying their search behaviours, and our understanding and hence the tools we develop needs to keep up with these social changes.

Following this line of argument, we can identify a further behavioural aspect of internet behaviour. How often have you, without real direction, stumbled across a whole load of material that you scan and then move on, only to find that a few weeks later you're reminded of it and want to look at it in detail? Rediscovering this material can be very difficult, since it's not entirely clear what it is you are searching for in the first place. Often the only way to do it is to return to the original activity and try to recreate the unstructured approach that led you to it the first time. It is the very nature of the internet that potentially unrelated information is encountered when searching for something entirely different[2]; it is the very nature of people that we are often distracted, entertained, amused, shocked or whatever by this, spend some time with it, then return to our previous activity. Simply finding things is not the issue; it's finding things that are directly relevant, without offering too much in the way of distractions.

This is an example of the wider issue of returning things that are of relevance to the user, whilst realising that the user works in many different situations and the notion of relevance varies from context to context. In particular, the context can vary between searching behaviour and browsing behaviour, often intermingled but offering different user experiences. Tools to support users therefore have to be able to make the transition seamlessly between these different modalities.

3 Supporting Users

The breakdown of internet interaction behaviour into three entities allows us to direct and focus our efforts on providing tool support for users. However, it is critical to realise that any tool that is developed must allow the free-from switching of context and behaviour modality without interruption. It is not good enough to simply support searching, for example, if it makes browsing difficult.

Search has been addressed by many people and web sites over the recent years, so that searching, per se, is not always the real issue. In fact, some would argue that searching on the internet is hardly a problem at all. There are a host of search engines, each with different characteristics, strengths and weaknesses, and for many people the internet is now the first choice of resource when it comes to looking something up. To use it effectively you need to understand something about the search engine, and to be aware that the veracity of information found cannot always be determined, yet it is still a major component of people's questing lives. Users are becoming familiar with the nuances of specific engines and are showing increasing ability to construct complex search queries to locate the information they want. Using '+' to force inclusion of terms, '-' to exclude them, and so on, are all examples of learned, targeted behaviours that allow users to access the full power of their search tools. The internet is the resource of choice for many, who access what they need through a combination of the available search tools, iterative refinement of their search terms, and an intuitive understanding of their own behaviour. However, this form of search is only keyword based; more complex semantic search systems are evolving and are the subject of related work.

3.1 Focus on Browsing

We have focussed on supporting browsing. Browsing the internet can be seen as a loosely directed traverse of a series of disconnected tree structures, with backtracking. Pages are nodes, with links as branches. Which link is taken is governed by whether the user finds it interesting enough to follow; if taken, it leads to a new page and potential new links being followed - alternatively, the path may be retraced back to a more interesting point, or that tree abandoned and a new one started. We employ intelligent modelling to determine the context of the internet interaction and hence provide some guidance through the multiplicity of options. Our aim is to give the user some guidance as to which parts of the tree are likely to be of interest to them, without cutting off any options.

The system is called 'Mitsukeru', Japanese for "to find out, locate". It uses an agent-based system to model the user's behaviour and determine the context of their interaction, and then looks ahead at the web pages linked to from the current page to determine their relevance to this particular interaction. The system is detailed below.

The system can be broken down into three parts; determining the current browsing context, determining the relevance of future pages based on the current context, and communicating this to the user. The system adheres to the principles identified earlier, namely allowing freeform switching of modalities between internet behaviours, and non-intrusiveness in that it operates seamlessly between the browser and the internet - all user interaction is still via the browser.

Mitsikeru incrementally builds a 'master' table for each browsing session based on word frequencies found in pages. This table consists of words and their frequencies. Very common words are discarded, and the table is added to for each page that is visited, and pruned to reflect the swaying of interest during a browsing session. This table therefore represents the current browsing context for this session; it may represent one or more browsing tasks.

The system uses a proxy to look ahead to the pages linked to the current page. which it then examines on behalf of the user. The proxy performs look-ahead and also parses the HTML to return the text of the page, removing tags and scripting information. The look-ahead therefore works for dynamically-produced pages, but not for image-based ones. Mitsikeru produces a 'page' table for each of the pages linked to in which the page is represented as a table of words and their frequencies. It uses these page tables to update a 'history' table, which is a list of all the words ever seen. We use these tables to determine relevance, using a Bayesian approach. There are many alternative approaches to determining relevance: we could have used entropy methods, or Chi-squared tests, or other latent semantic indexing approaches. The intention was to provide an algorithm that was relatively simple and fast to compute. Its benefits come from supporting the task effectively and presenting information efficiently. Different users will have different perceptions of relevance in any case, so any more complex measure is still doomed to psychological failure in terms of being more accurate. We are aiming to support and guide, not dictate and remove thought.

Words that are highly common (as defined by the history table) are likely to occur in the next page. Words that are not common, on the other hand, but which do occur on the next page, are of more interest. Of even greater interest are uncommon words that exist in both the current browsing context and the next page, as the chances of these occurring by chance are low. This approach to calculating relevance allows us more leeway in calculating the current browsing context, as we need not model each separate browsing task but can collate many tasks into a single representation.

3.2 Defining 'Interesting'

With n_{PT} as the number of words in the page table, and n_{MT} as the number of words in the master table, with $P(h)$ as the historic probability of a word, we can write the probability that the word occurs in the page table as

$$P_{PT} = 1 - (1 - P(h))^{n_{PT}}$$

The probability that the word occurs in the master table is

$$P_{MT} = 1 - (1 - P(h))^{n_{MT}}$$

The probability that the word exists in both the current and the master table is therefore

$$P(both) = P_{PT} \times P_{MT}$$

We now have probabilities that describe how common or not our word is in the both the current context and in the linked page under consideration, compared to the medium overall. We can write the *surprise factor* as:

$$P(\textit{one or other or both}) = P_{PT} + P_{MT} - (P_{PT} \times P_{MT})$$

Our definition of an interesting word is one that is neither a chance occurrence, nor very common, which we can write as follows:

$$P\ (\textit{keyword}|\textit{surprise factor}) = \frac{P_{PT} \times P_{MT}}{P_{PT} + P_{MT} - (P_{PT} \times P_{MT})}$$

Only interesting words are added to the master table. Interesting words lie towards the zero end of the spectrum. (In the implemented system, we chose words in the range 0.0006 to 0.1010). Note that this approach does not require us to filter the pages to remove HTML formatting, as this is treated as common and uninteresting. However, a stemming algorithm on the text would have increased the generalisation of the system without compromising the algorithm. Extensive use (and hence a sufficiently mature history and master table) obviates the need for this too.

We add the number of interesting words found on a page, and scale this by the number of words on the page, returning the result as a percentage, to arrive at our final relevance value. This favors shorter pages with more interesting words over longer pages, which is intuitively correct, though we recognise many alternatives are possible.

3.3 Information Presentation

Having calculated the relevance of the linked pages we have to present this information to the user. It is imperative that this is done in a non-intrusive manner [3], allowing the user to maintain their conventional browser usage, and not stopping them from switching behaviours. We achieve this though the proxy adding a layer of DHTML to the current page. Links that are dead are removed, whilst others are colour-coded according to their relevance. Strongly coloured links are directly relevant, whilst irrelevant links are closer to the standard text colour. This gives an immediate guide to the links that are most likely to be profitable followed, without cutting out any options. When the user hovers over a link, a summary of the page appears (title, initial sentence, main headings) and a relevance score is displayed in a small post-it style popup, achieved using DHTML. This allows the user to gain more information about the content of the page before actually deciding to visit it, and allows them to assess which of the likely candidates they should explore next much more rapidly.

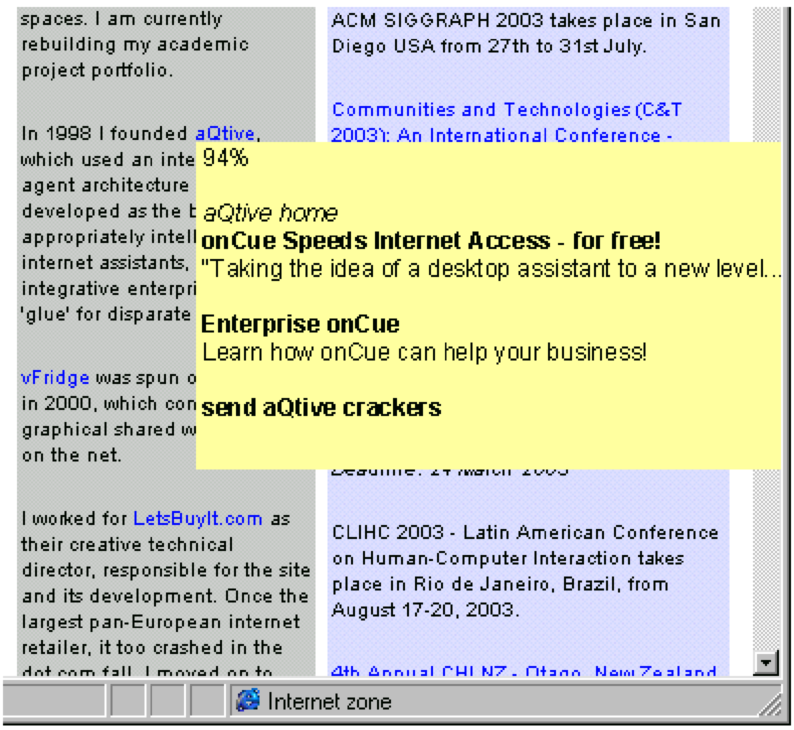

Figure 1: Hovering over a link brings up a summary of that linked page as well as a score that relates to its likely relevance to the current task

The system assists in search behaviour as well. Since browsing is done via a proxy server that acts as a cache for pages, we can bias any search towards pages that have been recently browsed. This means that pages that are an equally good keyword match but have been recently looked at are ranked much higher in the returned results. This allows us to more easily return to briefly seen information that we want to subsequently study in further detail.

4 Evaluation

The system has been designed to primarily address the browsing needs of users. As a task with flexible or ill-defined goals, it does not seem appropriate to provide quantitative measures showing, for example, reduced time to find a particular item, or other such clearly defined tasks. Instead, we have undertaken informal studies in which users have worked with the software and reported back their qualitative views. Users were drawn from a mixed population of science, engineering and humanities students, academics and the wider public. They ranged across the spectrum from hourly internet users to those who had only recently started using the internet. Their qualitative views can be summarised as follows:

- Mitsikeru is easy to learn; it has the same browser interface and interacts with web pages in the same manner as they are used to.
- The colouring of links is generally successful, with people tending to follow the strongly advised links most of the time. However, some users reported often wanting to see what was behind the other links to see how correct the system was.
- Presenting summary information about the next page was generally very well received, though the content and location of that information was sometimes criticised.
- Transitions between different browsing threads were not always carried out effectively by the system.
- The ability to inform the system as to its success or failure in recommending pages was requested by a number of users. This is an interesting point since it highlights that users are prepared to put in effort in helping the system become more accurate, and means that we do not have to solely rely on improved implicit techniques.
- Higher levels of satisfaction when browsing were reported.

Having a metaphor to understand how the system works was also mentioned by a number of people: they wanted to try and see why the recommendations were being made. The system as described is constructed from a number of independent but communicating modules: the pre-fetching proxy, the DHTML additive editor, the context modelling, the calculation of relevance, the page summary creation. Many of these modules act autonomously (without input from the user) and semi-intelligently (adapting their behaviour depending on the pages seen, providing enhanced information effectively 'hidden' from the user). We have found, through discussions with users, that they are best described as “agents”. Without getting in arguments as to definitions, users find that considering these processes as agents allows them to build up a more complete understanding of the system, and as a metaphor for user understanding, agents offer a positive contribution.

5 Best Practices

There are a number of best practices that this paper demonstrates. The key issue is that the system design goals were set by observing users in their interactions with the internet, developing an understanding of the types of activities they were involved in and then being able to categorise these. These categories of searching, browsing and monitoring then allowed us to identify areas that were less well supported, and focussed our approach to the systems design.

It is also best practice for supportive systems to be unobtrusive and subtle – these factors governed our approach to showing the additional information on the page through the dynamic rollovers as well as collecting the user information in the first place. We are looking for additional benefit without altering existing activities and strategies, and so must not be interrupting or annoying. In situations where we can gather the necessary information implicitly, without involving users in distracting dialogue, we believe that it is best practice to follow this approach.

Because the system supported deficiencies identified by observation and user involvement, it was easy to explain what was being supported; coupled with the strategy of inferring what we needed to know, the system was therefore easy to introduce to users and simple for them to understand what it was doing for them. Providing people with simple and compelling reasons to use a system is another example of good practice.

6 Conclusions

Tools such as Mitsikeru show us that by understanding some of the issues more clearly, it is not complex to provide more effective support for user behaviour. We have taken a principled approach to the development of the system. Observations of behaviour allow us to provide a useful categorisation of modalities of behaviour into three types, each distinctly different, though with very weak boundaries between them. This then allows us to focus on supporting one of these activities more effectively, without limiting the users freeform access. Using agent-based look-ahead and statistical measures of relevance, we are able to seamlessly interject additional information on the likelihood of linked pages being of interest without burdening the user with any additional work. Informal trials with users have shown a high degree of acceptance by them of the system, with reduced backtracking and a higher degree of satisfaction when browsing reported.

7 References

1. Beale, R, Wood, A, Drew, N S, & Hendley, R J (1995). HyperSpace: Web Browsing with Visualisation. Third International World-Wide Web Conference Poster Proceedings April 10-14 1995, Darmstadt Germany pp 21-25
2. Beale, R, Brayshaw, M.C., Hendley, R.J., Drew, N.S, Wood, A.M. (1995). Combining Agents and Visualisation for large information spaces. Proceedings of CKBS-SIG workshop on Intelligent Agents and the Next Information Revolution. Manchester..
3. A. Dix, J. Finlay, G. Abowd, R. Beale (1997). 2nd edition Human-Computer Interaction. Prentice-Hall, pp. 600.