Maintaining Consistency without Stagnation during Exploratory Search

Jaime Teevan
Microsoft Research
Redmond, WA 98052 USA
teevan@microsoft.com

ABSTRACT

Many real life search tasks are complex, multi-stepped processes. As people explore a search space, they want to find the most relevant information available to them while remaining oriented in the space they have already explored. This paper argues that during exploration it is important for search engines to present relevant information to their users in a way that maintains the users' existing contexts. This means that the most relevant results should not necessarily be ranked first, but rather ranked where users expect them to occur. The paper presents a model of what people remember about search results, and shows that it is possible to invisibly merge new information into previously viewed search result lists where information has been forgotten.

INTRODUCTION

Consider as an example of an exploratory search Connie's search for breast cancer treatments. Connie was recently diagnosed with breast cancer and wants to learn more about the available treatment options. For this reason, she runs a search for "breast cancer treatments". The result list returned to her is shown in Figure 1. Several results from the National Cancer Institute are listed first, followed by a result about alternative treatments, a link to About.com's page on treatments for breast cancer, and so on. The government pages appear too technical to interest Connie, and she is not generally interested in alternative treatments, so she skips over the first couple of results in the list and decides to follow the fourth link to an About.com page.

As Connie explores treatment options, it becomes possible for the search engine to identify results she may find more relevant to her search. Connie provides implicit feedback about what she considers relevant and irrelevant in the links she chooses to follow. She may also be willing to provide explicit feedback or query refinements because this topic is important to her. Further, her information need may evolve in predictable ways as she learns more about the topic, and new timely information about the latest treatments may become available as her search extends over time.

Although new, more relevant results can benefit Connie, naively re-ranking the search results she has already seen to place the better results first is not necessarily the best way to help satisfy her information need. Connie has developed expectations about what results the search result list for "breast cancer treatments" contains during her initial interactions with the list. If, for example, the About.com page she clicked on was no longer ranked about fourth in the list, she would have trouble returning to it.

This paper explores how consistency can be maintained in search result lists during long search sessions where new

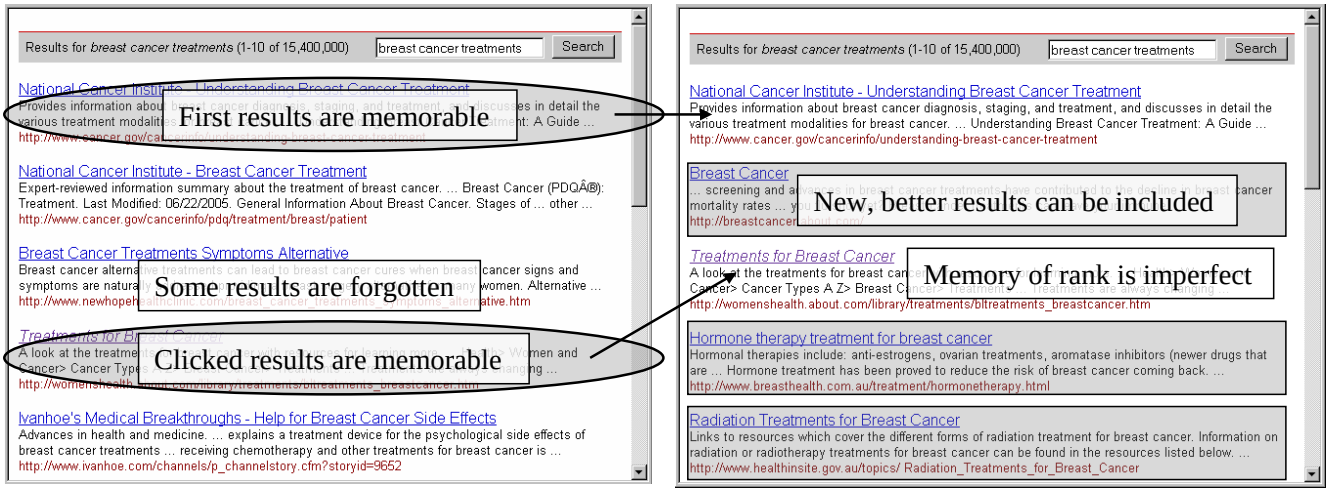

Figure 1. On the left is the search result list originally returned to Connie. On the right is a search result list that contains the results Connie remembers having seen before where she expects them that still includes new results.

information becomes available. It begins by demonstrating that changes to search result lists, even when beneficial, can cause disorientation. It then shows that changes can be made undisruptive by taking advantage of people's limited memories. Rather than keeping an entire result list static when a person returns to it, only that information that is remembered need be kept the same. New information can be snuck into the holes where results were forgotten.

Figure 1 illustrates the way that Connie's lapses in memory can be used to her advantage. Because she only remembers a few results in the original list returned for "breast cancer treatments", those results can be held constant while new more relevant results are added. The merged list is likely to look the same as the old list to her, despite containing new useful information. Sneaking new results into the result list can help Connie find new information while not interfering with her ability to re-find previously viewed information.

RESULT LIST CHANGES CAUSE PROBLEMS

Maintaining consistency is important because time and time again, changes to electronic information that should help the user get in the way. For example, dynamic menus were developed to help people access menu items faster by bubbling common items to the top of the menu. Rather than decreasing access time, research revealed that dynamic menus actually slow their users down because commonly sought items no longer appear where expected [4, 7].

Selberg and Etzioni [6] studied the rate of change of search result lists, and found result lists change rapidly. They noted that, "Unstable search engine results are counter-intuitive for the average user, leading to potential confusion and frustration when trying to reproduce the results of previous searches." Problems of this type appear in a study by White, Ruthven, and Jose [11]. In this study, the authors tried to help people search by giving them lists of relevant sentences that were dynamically re-ranked using implicit feedback gathered during the search. However, people did not enjoy the search experience as much or perform as well as they did when the sentence list was static.

The problems caused by changes to result lists were further explored in a study by Teevan, Adar, Jones and Potts [9]. They analyzed the queries issued by 114 people to the Yahoo search engine over a year, and found that when a previously clicked result changed position in the result list, users were less likely to re-click results. This suggests that changes to result ordering caused people to re-find less information and view more new information. A reduction in re-finding is not necessarily a bad thing if the new results are better than what was previously available. However, the authors observed that when the searcher clicked on a previously viewed result, the time it took to make the click was significantly longer if the result's rank had changed.

Because changes to result ordering slow re-finding, it is likely that highlighting previously viewed results, as is done by Google and A9, is not enough to support truly natural re-finding – not only because the result being re-found may have disappeared from the result list, but also because how it is re-found is not the same as how it was originally found. However, highlighting previously viewed results could be used to supplement the approach presented here.

Information management systems that permit a consistent interaction allow users to choose to interact with a cached version of their information space [2, 5]. Unfortunately, caching denies users the opportunity to discover new information. For example, Connie would not be able to revisit previously found information on breast cancer treatments while still learning about newly available treatments. People regularly find new information while re-finding. Although repeat clicks are common for repeat queries, 27% of all repeat searches involve the finding of new information as well [9].

MAKING CHANGES WITHOUT THE PROBLEMS

Fortunately, not all changes to search result lists result in a loss of context. It is possible for search tools to maintain the appearance of consistency while still presenting their users with the latest, most relevant information. This can be done by taking advantage of the fact that people do not remember all of the information they see. New information can be snuck into the holes where information has been forgotten. This section presents a model of what is memorable about search result lists, and shows how that model can be used to change unmemorable aspects of result lists to include new results without notice. A study is then presented that shows that sneaking in new results enables people to find new information quickly without destroying their interaction with previously viewed information.

Modeling What is Memorable

A study of 119 people was conducted to model what is memorable about search result lists [8]. In the study, participants were asked to interact naturally with a list of results for a self-generated query, and then later asked an hour later to recall the list without referring back to it. Two main factors emerged from the data as affecting how memorable a result was: where it was ranked and whether it was clicked. These factors were used to model which aspects of a search result list should be changed with care, and which could be changed freely. An intelligent merging algorithm was developed that merges new information into an existing result list by considering all permutations of old and new results and choosing the one with the highest benefit of both old and new information.

Model Allows New Information to be Snuck In

A second study was conducted with 165 different people to test the ability of the merge algorithm to hide change. Participants were asked to interact with a list of results, and a day later were asked to determine whether a follow-up result list was the same as or different from the original list.

Participants often noticed the follow-up list was different from the original list. When the follow-up list consisted of entirely new results, participants reported the list had

changed 81% of the time. When six random results were changed, the change was noticed 62% of the time, and when the clicked results were listed first and all other results were new, change was noticed 59% of the time. The differences between these cases are not significant.

However, when changes were made intelligently merging new information into the original list, the follow-up result list appeared static to participants. In this case, differences were observed only 19% of the time – less often, in fact, than the 31% of the time that a change was noticed when the follow-up list was unchanged. The disparity between the intelligent merging and the static list is not significant, but could possibly reflect the fact that changes to the result list made according to the model may create a list that looks more like the list the participant remembers than the actual original. The result lists from these cases were significantly more likely to be considered the same as the original list than any of the three cases with naïve change ($p<0.01$).

Merged Lists Support Finding and Consistency

Even though the merged list looks the same as the original list, the inclusion of new and better results can satisfy the user's information needs sooner. Usability improvements do not need to be noticed to benefit the user. A classic example is the Macintosh design for cascading submenus, where some flexibility in navigating to menu items is built into the menu design. The tolerance for small errors in navigation goes unnoticed by almost all users, but leads to fewer errors overall [10]. Similarly, a study of an improvement to cascading submenus showed all users performed better even though only three out of the 18 participants actually noticed the change [1].

To understand whether people were able to both maintain their context and find new information using the merged result lists, a subsequent study of 30 people performing re-finding and new-finding tasks was conducted. The study involved two parts: 1) an initial session where participants conducted initial finding tasks, and 2) a follow-up session where participants conducted finding and re-finding tasks.

When the list used for the follow-up session was the same as the list used during the initial session, participants were able to re-find information easily. However, they were unable to find new information. In contrast, when the follow-up list was comprised of entirely new information, participants were able to find new information easily, but could not complete the re-finding tasks. The intelligent merging performed closely to the best in both cases, allowing participants to find new information as if the list were entirely comprised of new results and re-find old information as if the list were unchanged. When the merging was done randomly, rather than according to the model, people performed comparatively worse.

Sneaking new information into the holes in a result list where people have forgotten results appears to be a good compromise for exploratory search. It allows consistency to be maintained during extended search sessions, but permits the inclusion of new information.

CONCLUSION

This paper investigated the importance of consistency during exploratory search. Changes to result lists were shown to create problems for people during the finding process. Although the ability to find new information may appear at odds with the maintenance of context, a solution was presented where new relevant results were not ranked first, but rather where the user expects them.

Looking forward, the effective management of changing information will be essential to successfully supporting complex finding behavior. The growing ease of electronic communication and collaboration, the rising availability of time dependent information, and the introduction of automated agents, suggest information is becoming ever more dynamic. Even traditionally static information like a directory listing on a personal computer has begun to become dynamic; Apple, for example, has introduced "smart folders" that base their content on queries and change as new information becomes available. As Levy [3] observed, "[P]art of the social and technical work in the decades ahead will be to figure out how to provide the appropriate measure of fixity in the digital domain." The solution presented here is a good first step towards that end.

REFERENCES

1. Ahlström, D. (2005). Modeling and improving selection in cascading pull-down menus using Fitts' law, the steering law and force fields. In *Proceedings of CHI '05*, 61-70.
2. Hayashi, K., Nomura, T., Hazama, T., Takeoka, M., Hashimoto, S., and Gudmundson, S. (1998). Temporally-threaded workspace: A model for providing activity-based perspectives on document spaces. In *Proceeding of HyperText '98*.
3. Levy, D. (1994). Fixed or fluid? Document stability and new media. In *Proceedings of European Conference on Hypertext*.
4. Mitchell, J. and Shneiderman, B. (1989). Dynamic versus static menus: An exploratory comparison. *ACM SIGCHI Bulletin*, 20(4): 33-37.
5. Rekimoto, J. (1999). Time-machine computing: A time-centric approach for the information environment. In *Proceedings of UIST '99*, 45-54.
6. Selberg, E. and Etzioni, O. (2000). On the instability of Web search engines. In *Proceedings of RIAO '00.*
7. Somberg, B. L. (1986). A Comparison of Rule-Based and Positionally Constant Arrangements of Computer Menu Items. In *Proceedings of CHI/GI '86*, 255-260.
8. Teevan, J. (2006). What People Remember About Search Result Lists. In *Proceedings of CHI '06*.
9. Teevan, J., Adar, E., Jones, R., and Potts, M. (2005). History repeats itself: Repeat queries in Yahoo's query logs. In *Proceedings of SIGIR '06,* 703-704.
10. Tognazzini (1999). A quiz designed to give you Fitts. http://www.asktog.com/columns/022DesignedToGiveFitts.html.
11. White, R., Ruthven, I., and Jose, J.M. (2002). Finding relevant documents using top ranking sentences: An evaluation of two alternative schemes. In *Proceedings of SIGIR '02*, 57-64.