

Web Browsing agents

SNU OOPSLA LAB.
서치영

Content

- Definition : What is the web browsing agent?
- Two kinds of the Browsing agent architecture
 - Server based approach
 - Client based approach
- Design principles for Web Browsing agents
- Letizia : An agent that assists Web Browsing
 - Autonomous Interface Agent
 - Interleaving browsing with automated search
 - Inference from the user's browsing behavior
 - Search strategies and recommendation process
- Conclusion

Question

What is the autonomous browsing agent ?

What is the web browsing agent ?

Definition : The agent that assists a user browsing the Web by analyzing a user's browsing behavior

The diagram illustrates the interaction between a user and a web page. On the left, a globe icon represents the 'Web'. On the right, a person at a computer represents the 'User'. In the center, an 'Agent' is shown as a small robot-like figure. A double-headed arrow connects the User and the Web. A speech bubble from the Agent says 'I recommend you this pages!'.

Two kinds of the agent architecture

- Server based approach (WebWatcher)

The diagram shows a central 'Agent server' in an oval. To its left is a starburst shape labeled 'Web'. To its right is a person at a computer labeled 'Client'. Two arrows, both labeled 'Browsing & search', point from the Agent server to the Web and from the Agent server to the Client.

A Web browsing agent is running as a separate server

Two kinds of the agent architecture

- Client based approach (Letizia)

The diagram shows a person at a computer labeled 'Client' with an 'Agent' icon on their screen. To the left is a starburst shape labeled 'Web'. Two arrows point from the Client system to the Web, representing the agent's interaction with the web.

A Web browsing agent is running in the client system with a user

Design principles for Web Browsing agents

- Suggest rather than act
: WBAgent only offers a suggestion that is "better than nothing"
- Take advantage of information the user gives the agent
: The actions taken by the user constitute information that the agent can use to infer the interest of the user
- Take advantage of the user's think time
: Running the agent autonomously while the user is thinking takes advantage of compute time otherwise be wasted.

Design principles for Web Browsing agents

- The user's attention may be time-shared
: A consequence of running an agent autonomously is that the agent cannot get the full attention of the user
- The Agent may have a different tradeoff between deliberation and action
: WBAgent will get more chances to improve its suggestion to the user as the user is browsing the Web

Letizia(Web Browsing Agent)

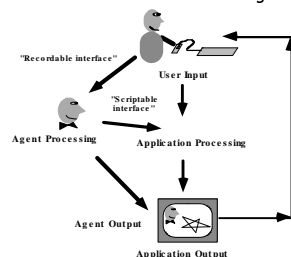
- By MIT Media Laboratory(1995-current)
- An autonomous interface agent for Web Browsing
- The behavior-based interface agent
- autonomously running in the client system
 - runs as a separate process
- Implementation
 - Language : Macintosh Common Lisp
 - Web Browser : Netscape
 - Communication between Lisp and Netscape : AppleEvent and AppleScript interprocess communication

Autonomous Interface agent (1)

- Interface agent
 - a software that actively assists a user in operating an interactive interface
 - makes changes to the objects which the user sees on the screen by observing many user inputs
 - The user may operate the interface regardless of the agent
 - But, if called upon, the agent may display suggestions or perform some actions on objects
 - Becoming more and more attractive due to the increasing complexity of user interfaces

Autonomous Interface agent (2)

- Structure of an interface agent application



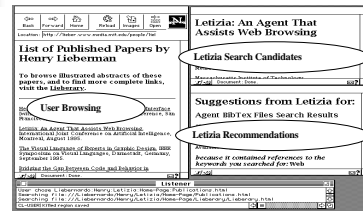
Autonomous Interface agent (3)

- Autonomous Agents
 - Software that operates in parallel with the user
 - may discover a condition that might interest the user and independently decide to notify the user.
 - Why autonomous agents ?
: An assistant may not be of much help if it needs very explicit instruction from the user
 - Allowing an interface agent to run in parallel with the user enables the user to delegate tasks to agent

Autonomous Interface agent (4)

- Autonomous Interface agents
 - Software that assists the user in operating an interface and operates in parallel with the user.
 - The user may or may not be aware of the agent's activities at any given moment.
 - Letizia : An autonomous interface agent for Web Browsing
 - Letizia presents its result to the user continuously
 - Letizia uses Netscape's interface to present its result , using an independent window

Letizia's default screen layout

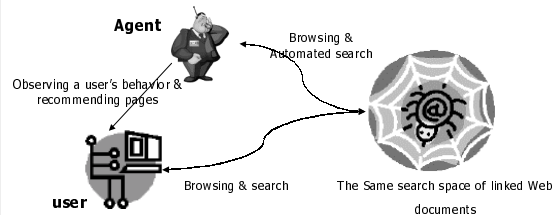


- The user may choose to continue browsing with either his own selected pages or Letizia's suggestions at any moment.
- Accepting a suggestion from Letizia simply consists of switching to a window that contains a page recommended by Letizia

Interleaving browsing with search(1)

- Letizia and the user both browse the same search space of linked Web documents
- The user can manually browse documents and conduct search without interruption from Letizia
- Letizia's role is merely to observe and make inference from observation of the user's actions
- In parallel with the user's browsing, Letizia conducts a resource-limited search to expect the future needs

Interleaving browsing with search(2)



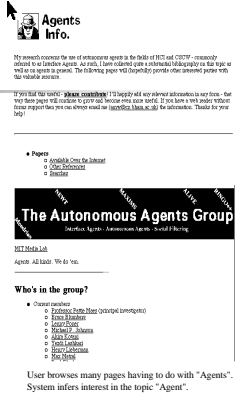
Inference from the user's behavior(1)

- The user's behaviors indicating his/her interests
 - Storing a reference to a document explicitly
 - The decision to follow further links
 - Entering the keyword in a page containing a search form
- The user's behaviors indicating his/her disinterests
 - Returning immediately without having stored the target document
 - Not following further links
- To compute the content of a document,
 - Not having natural language understanding
 - Use a simple keyword frequency measure, TFIDF (Term Frequency times Inverse Document Frequency)

Inference from the user's behavior(2)

- To Update a profile of the user's interests
 - step1: recording the URLs chosen by the user
 - step2: reading the pages
 - step3: analyzing the pages using a keyword-frequency measure (TFIDF)
- Letizia's judgement of interest in a document
 - Not trying to determine some measure of how interesting the document is in the abstract
 - Instead a preference ordering of interest among a set of links

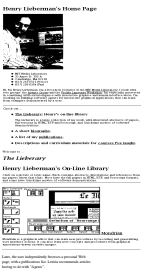
An example



The screenshot shows a page titled "Agents Info" with a search bar and a list of papers. The papers listed include "The Autonomous Agents Group" and "Who's in the group?".

- The user is interested in topics involving Agents
- The user may browse some general Agent Info pages
- or search for the word "Agent" in a search page
- So the agent can infer an interest in "Agents" from the browsing behavior

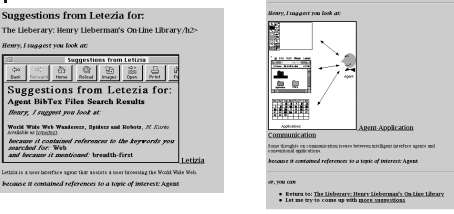
An example



The screenshot shows a page titled "Henry Lieberman's Home Page" with a list of publications. The publications listed include "The Librarian: Henry Lieberman's On-Line Library" and "The Librarian: Henry Lieberman's On-Line Library".

- At a later time, the user is browsing personal home pages
- A personal home page may contain a list of the publication
- As the user is browsing through some of the publications,
 - Letizia scans concurrently a list to find which ones may be relevant
 - Letizia suggests those papers dealing with "Agents"

An example



The screenshot shows a search results page with suggestions from Letizia. The suggestions include "Suggestions from Letizia for: The Librarian: Henry Lieberman's On-Line Library" and "Suggestions from Letizia for: Agent BibTex Files Search Results".

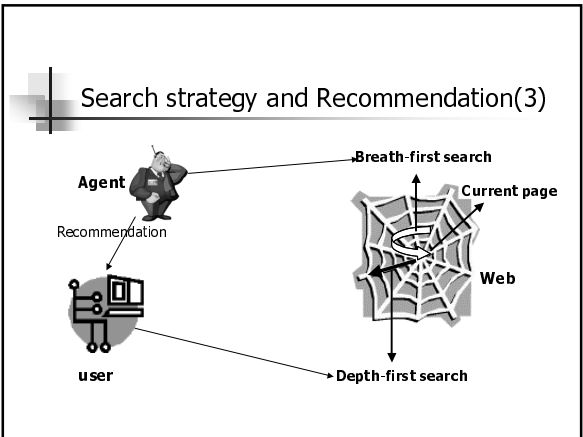
- Letizia can also explain why it has chosen that document

Search strategy and Recommendation(1)

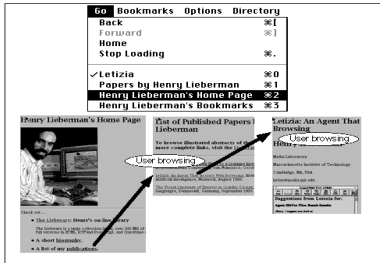
- The user generally moves "down" in the Web hierarchy
 - browsers tend to encourage a depth-first exploration of the Web space
 - Much information of interest to users is not typically located very far from the user's current position
- A breadth-first search done by Letizia
 - Solves problems emerged in depth-first search
- To control a breadth-first search by Letizia
 - put a resource limitation on search activity
 - ex) maximum number of accesses to non-local Web links per minute

Search strategy and Recommendation(2)

- To control a breadth-first search by Letizia (Cont'd)
 - Whenever the user switched from one Web page to the new page, Letizia's search is immediately refocused to the new page
 - The state of the search can be stored in a hash table indexed on the current page to continue if the user returns to the page
 - The user typically looks at a Web page for two or three minutes, then goes on
 - Thus the search tends to remain manageable

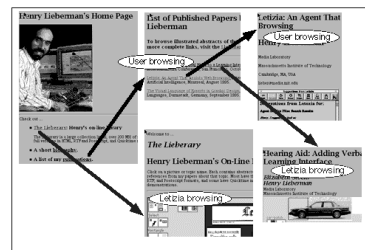


Search strategy and Recommendation(4)



Traditional browsing leads the user into doing a depth first search of the Web

Search strategy and Recommendation(5)



Letizia conducts a concurrent breadth first search rooted from the user's current position

Conclusion

- Definition of the Web browsing agent
- Two kinds of the agent architecture
 - Server based
 - Client based
- Design principles for Web Browsing agents
- Letizia : An agent that assists Web Browsing
 - Autonomous Interface agent for Web Browsing
 - Operating based on the user's behavior
 - Using TFIDF to analyze the content of page
 - Using a breadth-first search strategy