Supporting User Hypotheses in Problem Diagnosis on the Web and Elsewhere

Earl J. Wagner MIT Media Laboratory 20 Ames St Cambridge MA 02139 ewagner@media.mit.edu

ABSTRACT

People are performing increasingly complicated actions on the web, such as automated purchases involving multiple sites. Web services will only increase the complexity of these interactions. Things often go wrong, however, and it can be difficult to diagnose a problem in a complex process. Information must be integrated from multiple sites before relations among processes and data can be visualized and understood. Once the source of a problem has been diagnosed, it can be tedious to explain it to someone else, and difficult to review the steps of the diagnosis later.

We present a web interface agent, Woodstein, that monitors user actions on the web and retrieves related information to assemble an integrated view of a transaction. It manages user hypotheses during diagnosis by capturing users' judgments of the correctness of data and processes. These hypotheses can be shared with others, such as customer service representatives, or saved for later. We will see this feature in the context of diagnosing problems on the web, and discuss its broader applicability for debugging systems in general.

INTRODUCTION

The state of consumer interfaces for e-commerce is at a crossroads. Most Americans have access to the web and many are familiar with the convenience of late-night banking and the thrill of last-minute bidding. On the horizon is the advent of web services, currently being explored and developed for business-to-business interactions, but inevitably to have a direct influence on consumers' interfaces for purchases and online transactions. And yet most work in e-commerce interfaces for consumers—both in research and in the development of web sites—has been focused on what happens leading up to a transaction: how to provide appropriate and compelling recommendations, how often to send promotions and updates by email and other means, and so on. Interfaces for supporting customers after a transaction have received little attention and promising new trends like "customer selfHenry Lieberman MIT Media Laboratory 20 Ames St Cambridge MA 02139 lieber@media.mit.edu

service" are based on surprisingly old-fashioned technology like site-specific searching[6].

The issue of what happens when something goes wrong in e-commerce concerns vendors particularly because of the possibility of losing customers. A recent study found that after a bad customer service experience, 80% of customers would be less likely to buy from the online vendor again[3]. And customers are contacting support—nearly half (44%) of online buyers have made a support contact in the past six months[2].

Vendors will continue to work to reduce these numbers, both to build the value of their brands and limit the impact of customer service on their bottom lines. But we can expect the need for customer service to remain, just as we can expect the newest software with great features to always have a few bugs. The question then becomes to how to make the best of a bad situation.

Although the reasons for contacting customer service vary, some issues are more common than others. Just over a quarter of all contacts are due to a delayed delivery of a product or service and another quarter arise because of a billing or pricing issue[2]. Anyone who has made more than a few purchases online is familiar with these types of problems. Looking at your credit card transactions history, you wonder about that one order you never received. Or you purchased something then discovered that you received less, or paid more, than you expected. Diagnosing and resolving these problems and other can be an ordeal in itself. It helps to be organized and save order confirmation pages, as well as all the emails back-and-forth with a vendor-for every order you've ever placed. Then once you realize there's a problem with a particular order and try to figure out what went wrong, there's even more bookkeeping for documentation of conversations with customer service, product codes and serial numbers, and so on.

It is experiences like these that make consumers wary when initiating purchases in the first place. Indeed, the more complex a product or service is, the greater the quality of customer service will influence the decision[2], with financial services and airline and hotel reservations being the transactions of most concern. Importantly, these are also the transactions that are likely to span multiple sites, creating greater potential complications. Consider what happens when there's a miscommunication among multiple sites over the course of a transaction. Then it's the user's problem to resolve. The user has to go back and compare each step of what each vendor did with what it *should* have done. The user must assemble a complete history of the entire transaction, then return and both seek amends from the party at fault and resolve things with everyone else.

We see these as the areas that vendors should *really* be focusing on when developing new consumer-facing technologies. Web transactions are becoming increasingly complex and cannot be successfully handled with techniques developed back when we were told that "operators are standing by". Fortunately, we see an opportunity for software agents, working on the web, to provide advanced help for consumers in diagnosing and resolving their own problems, truly fulfilling the potential of customer self-service. An agent can retrieve information about an online transaction and visualize its entire history, even across multiple sites. In particular, it can show the flow of payments or items as they pass from one site to another. It can even be of help while diagnosing the source of a problem, by helping the user generate a hypothesis about the cause. All of this helps users "debug" the steps that a vendor took or that they took, or even just their own mental models of a process.

More importantly, we see this technique for annotating objects with user judgments as being more broadly applicable. This approach can help with current problems in ecommerce transactions, but also with monitoring processes on the web in general. It could help programmers during debugging and even help end-users diagnosing problems in systems in general.

In the rest of this paper we will discuss a software agent, Woodstein¹ with these features. Woodstein monitors user actions on the web and retrieves related information to assemble an integrated view of a transaction. We will see its data-history view, in which it presents the history of a transaction from the perspective of the transaction data such as prices and quantities of stock. We will also see the agent work at an even more abstract level, in supporting the annotation of objects in its data-history view, including transaction data, to help users manage their hypotheses while diagnosing a problem. We will see how these hypotheses can be shared as well as just saved for later reference. We will finish by sketching out the broader possibilities suggested by user annotations for problem diagnosis.

OVERVIEW OF WOODSTEIN

Woodstein is a software agent that works with a user's web browser to answer questions like "How did that data get that value?" "Why did that happen?" and "What's happening now?". It monitors a user's actions on the web to create a record of the user's overall process. For example, by watching the user browse an online retailer and add items to a shopping cart, it recognizes that the user is making a purchase. Later, when the user is looking at another page related to the process, Woodstein modifies the information in the page so that it can be directly inspected. Within the user's credit card transactions history page, a single charge can be examined. The history of the overall purchase process can be retrieved and reviewed, making it convenient to understand the context of the data, such as how the charge amount was computed.

Woodstein works by matching a user's actions to the steps of an abstract model for the process. Through this recognition, it knows to look for more information about the process on other web pages and web sites, even if the user never visited them. By watching the user select a credit card and shipper for a purchase, Woodstein knows to go to the sites of the bank and shipper to gather more information about the status of the purchase, including whether it has been paid for and delivered.

Woodstein collects and presents information about a user's data items and processes. A data item can be *simple* such as prices, addresses or dates, or it can be *composite* such as an entire transaction record or order. A process is either a user action, such as loading a page or clicking a link, or a web site reaction such as creating a new order. It is then able to explain the context and history of processes and data described in pages, such as how the items appeared in the shopping cart page. It answers questions about the history and current status of the overall process, as well as how data in the process was set.

Successfully diagnosing a problem requires an understanding of the causal relations within a system. Woodstein provides a "data-history" view to show the history of how a data item was computed and created. The user can revisit previous pages in its history within this automatically generated audit trail. For a purchase, the user can jump from the charge amount in the credit card transactions page to a saved copy of the order confirmation page in which the purchase price appears.

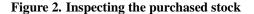
If the user feels that a process or data item looks incorrect, the judgment can be recorded through the object's contextsensitive menu. Woodstein then helps the user in diagnosing the source of the problem, even if it is just the user's incorrect understanding of the process—which is why we speak of an object as looking incorrect. Through the process of elimination, it makes further annotations and then suggests other objects to examine. With these "effective" annotations that trigger behavior by the agent, users save their intermediate judgments and have a record of their process of diagnosis.

¹Named after Bob Woodward and Carl Bernstein, the Washington Post reporters who uncovered the Watergate scandal. When their editor, Ben Bradlee, wanted to know what they had discovered, he'd stand at the door of his office and yell "Woodstein!' into the newsroom to call them in.

00		Corporate Empl	oyee Fina	ancial Se	rvices - Mozilla	a	C
6 0	6	http://127.0	0.0.1:8080	/www.sn-	afu.com/account	.html 🔽 🔍	. «. П
🟦 Home 🗎	Bookmarks						
SN-A	FU						
orporate Empl	loyee Financial S	ervices					
ock Options	Stock Purchase	Advice and Guidan	ice Mark	ket Data	Logout		
Yovod	vne Sto	ck Purchas	e				
		ch i urenus	•				
	·						La a se la se
Home Share	·	Order Entry	Statemer	nts Chi	ange Elections	ibrary & Forms	My Profile Help
	·	Order Entry	Statemen	nts Chi	ange Elections L	ibrary & Forms	My Profile Help
Home Share	Info Review C	Orders Order Entry		nts Chi	ange Elections L	ibrary & Forms	My Profile Help
Home Share	Info Review C		S	Order	Transaction Type		My Profile Help Share Delivery Method
Home Share Yoyodyne Reference	Employee -	- Review Order	s Order	Order	Transaction	Execution	Share Delivery
Home Share Yoyodyne Reference Number	Employee - Entry Date	- Review Order Share Quantity	s Order Status	Order Type	Transaction Type	Execution Date	Share Delivery Method
Home Share Yoyodyne Reference Number NX3333	Employee - Entry Date 08/29/2002	- Review Order Share Quantity 250	s Order Status EXE	Order Type N/A	Transaction Type TXFR	Execution Date 08/30/2002	Share Delivery Method transfer
Home Share Yoyodyne Reference Number NX3333 NX2222	Employee - Entry Date 08/29/2002 08/24/2001	- Review Order Share Quantity 250 406	S Order Status EXE EXE	Order Type N/A N/A	Transaction Type TXFR TXFR	Execution Date 08/30/2002 08/25/2001	Share Delivery Method transfer transfer
Home Share Yoyodyne Reference Number NX3333 NX2222	Employee - Entry Date 08/29/2002 08/24/2001	- Review Order Share Quantity 250 406	S Order Status EXE EXE	Order Type N/A N/A	Transaction Type TXFR TXFR	Execution Date 08/30/2002 08/25/2001	Share Delivery Method transfer transfer

Figure 1. Viewing a stock purchase transaction

Reference Number	Entry Date	Share Quantity	Order Status	Order Type	Transaction Type	Execution Date	Share Deliver Methor
NX3333	08/29/2002	250	EXE	N/A	TXFR	08/30/2002	mansfe
NX2222	08/24/2001	406	EXE	N/A	TXFR	08/25/2001	nsfe
NX1111	08/26/2000	418	EXE	N/A	TXFR	08/29/2000	sfe



AN EXAMPLE OF HOW WOODSTEIN MANAGES A USER'S HYPOTHESIS

Woodstein's support for visualizing a user's actions and managing user hypotheses can be seen with an example. Consider an employee at Yoyodyne who is enrolled in his company's stock purchase plan. Each pay period Yoyodyne sets aside a portion of his paycheck and, once a year, uses this money to purchase a block of Yoyodyne shares with its broker, SN-AFU. He has set up his account at SN-AFU to automatically transfer the shares to his broker, Sellwell.

The employee is browsing on the web and decides to review Yoyodyne's share purchases at SN-AFU. He notices that the number of shares most recently purchased seems lower than usual at 250, rather than over 400 (Figure 1).

The employee wants to interact with the shares directly to see their history. When analyzing this page, Woodstein added its logo that overlays the bottom right of the page. The employee turns on Woodstein's inspector by clicking on the logo. Woodstein converts all of the objects it recognized while analyzing the page to buttons (Figure 2). When the user moves the mouse into a button, Woodstein darkens the button and updates the browser window's status bar to indicate what will clicking will do. This explanation also provides the actual name of the button's process or data item.

The employee wants to know how the number of shares purchased was set and presses down on its button, revealing a menu. He selects "How was this set?" (Figure 3).

Deliver	Execution Date	Transaction Type	Order Type	Order Status	Share Quantity	Entry Date	Reference Number
mansfe	08/30/2002	TXFR	N/A	EXE	250	08/29/2002	NX3333
nsfer	08/25/2001	TXFR	N/A	EXE	Why was this set?	08/24/2001	NX2222
sfer	08/29/2000	TXFR	N/A	EXE	How was this set?	08/26/2000	NX1111
STREET, BARRIER	08/29/2000	TXFR	N/A	EXE		08/26/2000	NX1111

Figure 3. Asking how the stock was purchased

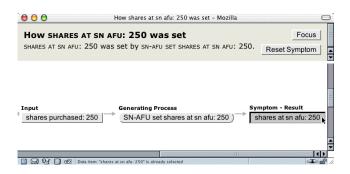


Figure 4. Viewing how the stock was purchased

Woodstein opens a pop-up window showing the history of the shares with English descriptions of the processes and data involved (Figure 4). This window is the "data-history" view and it shows how the data item "shares at SN-AFU" was set. The top frame, in grey, answers the employee's question by indicating that the shares resulted from the process "SN-AFU set shares at SN-AFU". Each data item is set as the result of a process. Processes take data items as inputs and set a data item as the result. The bottom frame shows how the process took the number of shares purchased as its input and set the number of shares at SN-AFU as its output.

Woodstein created this record by matching the Yoyodyne's original concrete steps in exercising the employee's stock purchase plan with its abstract process model for the action. After Yoyodyne purchased the stock at SN-AFU, Woodstein knew to look at both web sites and match the stock transaction. In this view, Woodstein presents the information it gathered from the sites involved in the purchase, revealing the first few steps back of the history of the shares.

Processes and data items are both presented as buttons in Woodstein's inspection mode, but process buttons are rounded while data buttons are rectangular². All buttons for the same data item or process are equivalent across different views, so interacting with an object's button in the page is the same as interacting with it in the data-history view. Woodstein provides multiple views to show the different relations among the processes and data. For instance, a page shows the original context of processes and data items, but the data-history view shows how they are causally related. Each data item is set as the result of a process. Processes take data items as inputs and set a data item as the result.

²As is standard in data-flow diagrams.



Figure 5: Noting that the quantity of stock purchased looks wrong

Input shares purchased: 250	Generating Process SN-AFU set shares at sn afu: 250	Symptom - Result shares at sn afu: 250
III III OF Click to select proc	ess: "SN-AFU set shares at sn afu: 250"	

Figure 6: Viewing how the quantity of stock purchased was wrong

Returning to the scenario, the employee thinks that "shares at SN-AFU" looks incorrect. He presses down on its button in the data-history view and selects "This looks Wrong" (Figure 5). The button turns red to indicate it has been annotated as looking incorrect. Woodstein then marks the objects used to compute the data item, both the process that created it, "SN-AFU set shares at SN-AFU", and that process' input, "shares purchased", yellow to indicate that they may be incorrect (Figure 6).

In response to the employee noting the incorrectness of the data item, Woodstein opens a pop-up window to guide him through the rest of the process of diagnosing the problem (Figure 7). It put the objects it marked on the list of objects for him to examine.

In addition to automatically identifying objects for the user to examine next, Woodstein also performs the process of elimination to help the user identify the source of the problem. If the output of a process looks wrong, but all of the inputs look right then something must have gone wrong with the process. Similarly, if the output looks wrong but the process itself and all but one if its inputs look right, then the problem must reside with the remaining input.

The next step is to determine the correctness of the process that set the shares, "SN-AFU set shares at SN-AFU". Moving the mouse over the button for the process updates the debugging trail window which provides the employee with some guidance. If the problem looks like it started before the process, the process should be marked as correct. The employee notes that the input share quantity was also 250, so he marks the process as correct (Figure 8). Woodstein, in turn, narrows the problem to the input and updates the data-history view with both new annotations (Figure 9).

The employee still doesn't know why so few shares were purchased, though, so he continues. He clicks on the triangle

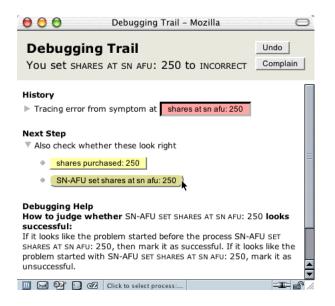


Figure 7. Viewing the record of the diagnosis



Figure 8: Noting that the stock purchase looks successful



Figure 9: Viewing how the quantity of stock to purchase looks wrong

Process Input Yoyodyne set shares purchased: 250	nerate hares
OPP O C Process: Yovodyne set shares purchased: 250" is already selected	

Figure 10: Viewing how the number of stock to purchase was computed

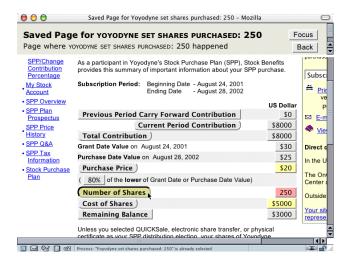


Figure 11: Viewing the saved page for the computation of the number of stock to purchase

next to the "shares purchased" data item to open its history, scrolls the history into view, and clicks on the the process that set the data item (Figure 10).

The process that resulted in shares being purchased was "Yoyodyne set shares purchased". Clicking on a process or data item causes it to become selected and its button to appear pressed in. The shares at SN-AFU were previously selected because the employee asked how they were set. Now the process for Yoyodyne setting shares purchased is selected. When an object is selected, Woodstein opens its "savedpage" view with page it saved for the object. The page is either the page the user interacted with directly, or a page the agent retrieved with the first appearance of the data item or a description of the process. In this case, the view features the saved copy of a retrieved page with the number of shares Yoyodyne intended to purchase (Figure 11). It is accessible through the "Number of Shares" label, which stands for the process that set the number of shares.

As in other views Woodstein presents the data items and processes it is tracking as possibly annotated buttons within the saved page. Just as in any other of Woodstein's views, the user can interact with any of these buttons to access the history of his data and the processes they were involved in.

The employee looks in the data-history view and sees the inputs to Yoyodyne setting the number of shares purchased. Yoyodyne used the share purchase budget and the share price in computing the number of shares. Within the saved page, he can see some of the history for the share price. Yoyodyne started with two prices, the price at the beginning of the period and the price at the end of the period. Then it took 80% of the lower of the value, which resulted in the share price of \$20. So far, that looks right, and the prices of the shares themselves look right to the employee so he marks them. He looks over the rest of the saved page and notices something unusual. His total contribution this year was \$8000, which looks correct, but for some reason only \$5000 was used to

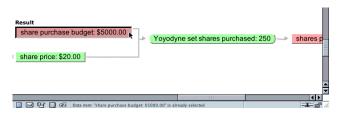


Figure 12: Viewing how the budget for the stock to purchase looks wrong

total contribution: \$8000.00	Yoyodyne set share purchase budget: S	5000.00) share Result Share
III 🔄 📴 🚺 🗗 Click to select data item: "Ir	rs limit: \$5000.00"	

Figure 13: Viewing how the IRS limit limited the share purchase budget

purchase shares. It looks like this may be the problem, so he goes back to the data-history view to see the history of the budget, and how it was computed using his total contribution. He clicks open the share purchase budget data item (Figure 13).

He sees that Yoyodyne set the budget, and this process took his total contribution and an IRS limit as inputs. It looks like limit may be the problem. He clicks on it to select it, opening the saved page that explains it in more detail (Figure 14).

It looks like this is in fact the source of the problem he ran into. This obscure and newly-introduced policy limited the amount that could be spent on his stock. With Woodstein, the employee was able to easily diagnose this problem and see that it was actually a problem with his own understanding of the stock purchase plan policies. He was able to see the history of how SN-AFU's transfer process and Yoyodyne's purchase process interacted with his own data to create the result on his SN-AFU account page. By tracing the history of the stock through the data-history view, he avoided having to look up the history of these processes on each individual site's pages. Tracing back into the history enabled him to see the exact policy that caused the unexpected result and, with the saved page view, he was able to see the explanation of the policy on Yoyodyne's site buried deep within its help pages.

The employee is happy to have identified the source of his problem. Now he wants to find out if there's an alternate program he can enroll without the restriction. He goes to the debugging-trail view and clicks on the "Complain" button to generate an email (Figure 15).

Within the email's user-editable area, he asks for more information. The CSR who receives this will be able to understand the exact context of the request based on the path the employee took.

00	Saved Page for irs limit – Mozilla	0
-	e for IRS LIMIT: \$5000.00 Focus with IRS LIMIT: \$5000.00 matched Back	4
	stock. The remaining \$30 is added to your payroll deductions during the next subscription period and used to purchase stock later. If you decide not to continue your participation, the balance is refunded to you, without interest.	
	Your shares will be electronically deposited into your SN-AFU account two business days after the end of the subscription period.	
	Note: The IRS now limits the number of shares that may be purchased in a calendar year through the Stock Purchase Plan to a market value of \$5000	
	Eligibility	
	You are eligible to participate in SPP if you:	
	 Are an employee of Yoyodyne or a participating subsidiary; and Regularly work 20 hours or more per week and at least five months per year; and you are employed as of the last day in 	4
I 🖂 🔮 🔲 dz	Data item: "irs limit: \$5000.00" is already selected	⊫ _^ //

Figure 14. Viewing the saved page for the IRS limit

subject. [Problem with the data: "irs limit: \$5000.00" which looks Wrong]	
I've noticed there's a problem with the data: "irs limit: \$5000.00" which looks Wrong at: www.yoyodyne.com/policy.html	•
Here's how I found out: I started with the data: "shares at sn afu: 250" which looks Wrong at: www.sn.afu.com/account.html	
then I saw the process: "SN-AFU set shares at sn afu: 250" which looks Successful at: www.sn-afu.com/account.html	
then I saw the process: "Yoyodyne set shares purchased: 250" which looks Successful at: www.yoyodyne.com/account.html	
then I saw the data: "share price: \$20.00" which looks Right at: www.yoyodyne.com/account.html	
then I saw the process: "Yoyodyne set shares purchased: 250" which looks Successful at: www.yoyodyne.com/account.html	
then I saw the data: "share price: \$20.00" which looks Right at: www.yoyodyne.com/account.html	
then I saw the data: "irs limit: \$5000.00" which looks Wrong at:	•

Figure 15. Complaining about the IRS limit

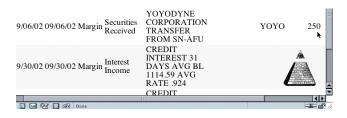


Figure 16. Viewing a stock transfer transaction

00	How shares at sellwell: 250 was set - Mozilla	0
How SHARES	AT SELLWELL: 250 was set	Focus
SHARES AT SELLWEL SELLWELL: 250.	LL: 250 was set by SELLWELL SET SHARES AT	Back
SELLWELL. 250.		Reset Symptom
Input	Generating Process	Symptom - Result
Input shares at sn afu: 2		Symptom - Result shares at sellwell: 250

Figure 17. Viewing how stock was transferred

We can see how beneficial Woodstein has been by looking at the steps the user would have to go through to find the same information without Woodstein. He would have seen the problem symptom at Yoyodyne's broker, SN-AFU. The next step would have been to visit Yoyodyne's internal pages. He would log in, find the section for his stock purchase plan account, and load the purchase history page showing how the stock was purchased, if it was available at all. At that point he would have to go back and carefully read the help for the stock purchase plan in order to find the single sentence describing the IRS limit. Alternately, he could call Yoyodyne's internal support. After waiting on hold and providing information identifying himself, he would talk with a CSR and, after both have traced through the process of the share purchase, he would eventually learn about the limit.

We can see how Woodstein's records of a user's exploration and diagnosis are helpful by continuing with the example after time has elapsed. A few months later, the employee is browsing the web and loads his transaction history for his broker, Sellwell. He sees the automatic transfer from SN-AFU and remembers that there was something unusual about it. He now wants to be reminded what the problem was (Figure 16).

He inspects the transaction, loads the data-history view and sees that he had inspected the transferred stock when it was still at SN-AFU (Figure 17).

In this example, we have seen how Woodstein presents the complete history of the purchased stock, even when it appears on multiple sites. Furthermore, we saw how it records the user's interaction with Woodstein's history record itself. The user can make use of the interaction record to, for instance, familiarize someone else with the identified source a problem. Later, Woodstein supports reviewing interaction history when other pages related to the process are visited. From user testing we have found that even first-time users of Woodstein can perform a simple diagnosis like we saw in the beginning of this example, in which the history of some data is accessed and the user traces back to a particular policy that affected it, in about 5 minutes.

EVALUATION

We tested the Woodstein's integrated view and complaint generator with 16 subjects [12]. We hypothesized that subjects who used one of Woodstein's views would be more successful in diagnosis and diagnose problems more rapidly. The eight subjects in the control group diagnosed problems spread across multiple pages of multiple web sites, then simply complained about the relevant data item or process. The eight subjects in the experimental group used one of Woodstein's views, its process-history view, to see an overall history of the steps of the process in identifying the data item or process to complain about. Woodstein's process-history view differs from its data-history view, presented earlier, by showing the entire process spanning multiple sites as a process tree much like the tree-view of a file browser.

Participants were told how the use the system and saw it

demonstrated. This took 5 minutes for control group and 20 minutes for the experimental group. They then spent about 10 minutes taking a "quiz" consisting of using the agent to select a particular data item and complain about it. For the experimental group, this required accessing a saved page through one of Woostein's history views. Participants then had 20 minutes to solve a "test" problem involving a simulated user's data and an organization's policies. In this problem, each participant played the role of a student who is unable to graduate from his educational institution and attempts to identify the exact reason and policy involved. The experimental group used one of Woodstein's history views, while the control group just used pages on the web site.

All participants in the experimental group of the user study were successful in diagnosing the test problem, taking an average of 5 minutes. Only two participants in the control group were successful, requiring an average of sixteen minutes. In addition, at the end of the experiment session, we asked participants in the experimental group to rate different aspects of their experiences with the agent and whether they'd use it for problems they might encounter. Interestingly, half would have used the agent to diagnose a problem if they knew it were to take longer than 5 minutes on the phone. Of the remaining four, three would use the agent if they knew a phone resolution would take 15 minutes.

Reflecting the difficulty of managing information about credit card purchases, half of the participants in the experimental group said they "strongly agree"³ that they'd like to have an agent like this one to help in visualizing and managing their credit card transactions. Two said they "agree" they'd like to have it, and the remaining two participants discussed the weaknesses they perceived in the agent later in a free-from section. One noted the potential for privacy problems and the other preferred a more interactive process for complaining. Though a CSR is interactive, no CSR is able to provide all perspectives of a transaction involving multiple vendors, as Woodstein can. The agent could be extended to support an interaction with a CSR, however.

In fact, one of the original motivations for Woodstein is the state of current support involving limited, text-based media including phone conversations and email. Even a perfect interaction with customer service by phone still lacks the benefits of a web-based interface. In order to resolve a problem, a customer must be sure to call during the hours support is offered, and have a block of uninterrupted time. The web, on the other hand, is always available and if some information is not immediately at hand, a decision or investigation can be postponed. Furthermore, it is often easier to understand complex information when it is presented graphically[7]. All of the advantages of a slick web site's presentation are lost when the customer has to hear his options or receive instructions over the phone. In fact, all of the advantages of a digital format are lost. With both the user and a CSR using an agent like Woodstein, both could highlight and talk about

³Participants expressed their level of agreement with a 7-point Likert scale: 1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=neutral, 5=somewhat agree, 6=agree, 7=strongly disagree

data items and processes, and even hypothetical possibiliites and future events.

MANAGING HYPOTHESES WHEN DIAGNOSING PROB-LEMS IN SYSTEMS

We see the possibility of user annotation as more broadly applicable beyond just e-commerce on the web. Computer users interact with systems everyday and often run into problems. When a problem is repeatable or particularly significant, a user may send a bug report in an informal way. Some systems, like Bugzilla[8], have been developed for automatically managing users' bug reports. Further, some applications, like Mozilla[9], "close the loop" and automatically send a bug report when erroneous behavior is detected, such as when the program crashes. These applications often allow users to provide some free-form information about what they were doing when the problem occured, but they don't support user diagnosis in general. Unlike a developer, however, a user is in the perfect position to diagnose hard-tofind bugs involving configuration details. Further, we suspect that many users would be interested taking a few moments to diagnose a configuration problem with their operating system then have to reinstall the entire system and all of their applications. In fact, often users have to do exactly that and reassemble the history of what they installed to identify a conflict. Of course, perfectly developed software with no bugs would be ideal, but other factors govern the adoption of new software. Regardless, a higher-level way of managing hypotheses during diagnosis would be helpful.

RELATED WORK

A companion paper[13] discusses Woodstein's basic behavior without object annotation and hypothesis management. It explains more fully how Woodstein tracks user actions and it presents Woodstein's other views.

No existing system that we are aware of directly attacks the problem of end user debugging of electronic commerce transactions or Web interactions.

Program Slices

A program slice is akin to the notion in business of an "audit trail". Audit trails are used to track the history of a record and show all of the processing it has undergone. Program slicing in particular is a software engineering technique for focusing only on the parts of a program that affect the value of a particular variable[14]. It is helpful for debugging, when a programmer knows a variable has the wrong value and wants to know how it was computed.

Program slice tools typically highlight the lines of code, modules or files in a slice[1]. This is useful for programmers, for whom the source code is the primary representation of the program. Within the domain of web actions, however, we don't expect the abstract models to be particularly meaningful to end-users. Rather than presenting the abstract description of the process, Woodstein generates explanations of the process' actual concrete execution.

Some tools present slices via control-flow graphs or program

dependency graphs[5]. Woodstein presents the program dependencies in the data history view, and the program execution tree in the process history view.

End-User Debugging

Some researchers have focused on how to better support endusers and novice programmers in debugging. In what she calls "end-user software engineering", Margaret Burnett in particular has focused on creating visualizations to support end-users in debugging spreadsheets [10].

Capturing User Annotations

Little research has focused on capturing user annotations, and less still on what we call "effective" annotations with a semantic meaning for the system that manages the annotated objects.

Trellis is a system that supports user annotation of objects in an applications interface[4]. Annotations in Trellis are typically used to indicate the original context of some information, allowing users to refer back to this context.

Third Voice is a system for managing user annotations of web pages, allowing users to see pages on the web overlayed with the comments of other users[11].

CONCLUSION

We have presented Woodstein, a web interface agent for enabling end-users to visualize and understand their actions on the web. Further, Woodstein helps users manage their judgements of the correctness of their data and processes and diagnose the sources of problems they run into. We saw an example of how this record of a user's judgements can be useful both to share with others, including customer service, as well as for future reference.

ACKNOWLEDGEMENTS

Thanks to Marc Millier for his help in developing the example scenario, Chris Laux for help in refining Woodstein's interface and Mary Jane for inspiration.

REFERENCES

- Thomas Ball and Stephen G. Eick. Visualizing program slices. In *IEEE/CS Symposium on Visual Languages*, pages 288–295, 1994.
- 2. David Daniels, Corina Matiesanu, and David Schatsky. Jupiter Consumer Survey Report: The State of Customer Service 2003. Jupiter Research, 2003.
- 3. David Daniels and David Schatsky. *Quantifying the Cost of Poor Service: Investing in Customer Service to Defend Revenues.* Jupiter Research, April 2002.
- 4. Yolanda Gil and Varun Ratnakar. Trellis: An interactive tool for capturing information analysis and decision making. In *Lecture Notes in Computer Science 2473*. Springer-Verlag, 2002.
- 5. Tommy Hoffner, Mariam Kamkar, and Peter Fritzson. Evaluation of program slicing tools. In *Automated and Algorithmic Debugging*, pages 51–69, 1995.
- 6. Esteban Kolsky. *The Six Steps for Web Self-Service in Customer Service*. Gartner, Inc., March 2002.
- 7. Richard Mayer. *Multimedia Learning*. Cambridge University Press, 2001.
- 8. The Mozilla Organization. Bugzilla: The mozilla bug database.
- 9. The Mozilla Organization. Mozilla.
- J. Ruthruff, E. Creswick, M. Burnett, C. Cook, S. Prabhakararao, M. Fisher II, and M. Main. End-user software visualizations for fault localization. In ACM Symposium on Software Visualization, 2003.
- 11. Third Voice.
- Earl J. Wagner. Woodstein: A web interface agent for debugging e-commerce. Master's thesis, MIT Media Laboratory, 2003.
- 13. Earl J. Wagner and Henry Lieberman. Understanding your actions on the web. In *Submitted to Proceedings of CHI'04*, 2003.
- 14. Mark Weiser. Program slicing. In *Proceedings of the Fifth International Conference on Software Engineering*, pages 439–449, New York, 1981. IEEE.