



Special points of interest:

- Gary Wolff signs the praises of Opera
- Kurt Keller continues about regular expressions
- Barbara Manning reviews *The Elements of User Experience*
- David Parry discusses word processing and translation
- Ann Colville opens the (refrigerator) door on a scary future

It's Not Over till the Fat Lady Sings *The New Opera 7.01 for Windows*

-- "the fastest, smallest, most full-featured desktop browser on the planet" --

by Gary Wolff

Well folks, this is one of those rare cases where, indeed, less is more.

The Norwegian company **Opera Software ASA** has just edged competition up a notch by releasing on January 28 its latest innovative tool for web surfing and I must say, I'm pretty impressed. Opera is now the third most popular browser on the web, and growing fast: around one million new users are downloading and installing Opera each month.

The Java-less installation program which I downloaded takes up less than 3.5 megabytes of memory. In addition to not taking up much memory or space on your computer, one of Opera's main appeals, I think, is its speed. Because it's small, and because it's written to keep things simple, most web pages load very quickly. For example, while an information-packed newspaper web site might take minutes to load in Microsoft Corp.'s Internet Explorer (IE), Opera

zaps it onto your screen in what seems like only seconds (independent browser performance tests conducted by **GreyMagic** can be viewed at: <http://www.greymagic.com/dagon/results.html>).

COOL NEW STUFF

Opera also has more, and better, features than its heavyweight competitor, Internet Explorer. For example, take the "Forward" button on

(Continued on page 3)

Got Computer Problems?

Join the Club!

The Tokyo PC Users Group is Japan's largest English speaking personal computer club, with a particular interest in bilingual computing.

Users of any platform or OS, of any ability, from occasional web surfers to genuine geeks and gurus should check out our on-line newsgroups and monthly meetings. We all have techniques to learn and tricks to share.

Inside this issue:

Opera 7.01/Gary Wolff	1
RegEx Part 2/Kurt Keller	8
Book Review/B. Manning	13
Ionic Column/D. Parry	15
Blewtooth Blues/A. Colville	18
Babitoon/A. Colville	20

About the Tokyo PC Users Group

www.tokyopc.org

Here you can find out about upcoming meetings and events, catch up on past meetings you may have missed, browse the newsletter online, and find out about all of the membership services.

(Contributors are always welcome. Mail your commons, suggestions, or interest in helping out to:

tpc-webmasters@tokyopc.org.

TPC Mailing Lists

The group provides mailing lists that mirror each of the newsgroups. Send a blank message to: info-lists@tokyopc.org to receive a catalog of the current lists and complete information on how to subscribe.

TPC Newsgroups

The group runs its own private news server that you can access over the internet from anywhere in the world. There are newsgroups for everything from help requests and Windows troubleshooting to non-computer topics like movies and dining around Tokyo. Set up your newsreader to point to news.tokyopc.org or visit <http://www.tokyopc.org/tpc/newsgroups.html> for a "no-setup" way of trying out the service.

Snail Mail Correspondence

Tokyo PC Users Group, P.O. Box 103, Shibuya-ku, Tokyo 150-8691, Japan

E-mail Correspondence

Mail: president@tokyopc.org, or any of the listed officers with questions or suggestions for the TPC.

Membership

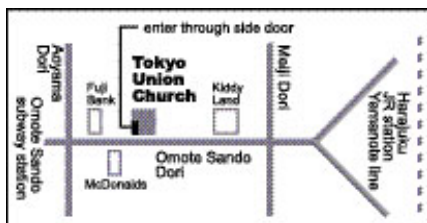
Individual: ¥10,000 per year. Includes newsletter, admission to regular monthly meetings, priority registration to SIGs and workshops, and other benefits.

Corporate: ¥25,000 per year. Full details:

www.tokyopc.org/tpc/join_tpc.html

Monthly Meetings

Meetings are the first Thursday of the month (2nd Thursday in case of Holiday conflict) in the basement of the Tokyo Union Church. Doors open at 6:30pm and the meeting begins at 7:00pm. Fee for non-members is ¥1000. See the map below.



From Omotesando station, take exit A1. When you reach the street, with the Hanae Mori Building on your left, continue downhill to the Tokyo Union Church on the corner of the third side street. Turn left at the church and use the side entrance inside the carport. Take the stairs down to the basement to the meeting room. Note that the TPC is not affiliated with any religious group. (Most of us head down the road to Shakey's for beer and pizza after the meetings.)

Executive Committee

President: Barbara Manning
barbara.manning@tokyopc.org

Vice President & Reviews Editor:
S. Patrick Eaton
Spe@l-sako.com

Treasurer: Anthony Whitman
03-3770-1513
tonyw@gol.com

Editor & Past President: Mike Lloret
090-9380-3508
lloret@gol.com

Web Publisher: Kevin Ryan
ryan@kevinryan.com

Publicity: Justin Gardiner
justin@fka.att.ne.jp

Corporate Liaison: Andrew Shuttleworth
andrew.shuttleworth@tokyopc.org

Disk Librarian: Mike Kato
mikekato@gol.com

APCUG Rep: Ken Cotton
(03) 3984-9600/4487 (tel/fax)
kc@gol.com

Programming SIG: Robert Altena
Robert@RobertAltena.com

Promotions: Sam Julien
sjulien@jp.psi.net

Webmaster Committee: Sajjad Zaidi
sajjad@vgkk.com

TPC Exec Meetings are held on the Monday following the General Meeting. All are welcome to attend.

Bank Details

Sumitomo Mitsui BC (Sakura) Bank,
Akasaka Branch #825
Regular Account #7174919
Name: Tokyo IBM PC Users Group

(Continued from page 1)

your usual browser. In other browsers it will only take you back to a page you've already visited, and then left. Right? Opera, however, instead tries to figure out from the page you're visiting where you'll go next, so that by clicking on the forward button you'll go there automatically. The forward button will change into a double-arrow "Fast Forward" button on pages where this is possible.

It works like this: Let's say you're using the **Google** search engine to look for the word "Lloret" on the Tokyo PC web site. After you type in site:www.tokyopc.org Lloret, the first page of matches appears, but to reach the following pages of over 100 matches, you'd have to keep scrolling down to the bottom of every Google page to find the "Next" link. Not with Opera: You simply click on the fast forward button which will take you to the next page of matches, and then to the next, and so on. Simple, elegant, and useful. It works equally as well when perusing multi-page online news articles.

If you find moving the mouse to the "forward" button to be a hassle, try this: Just press the "x" key and you'll automatically go to the next page of Google's Lloret matches. Or press the "z" key and you'll go back to the last page. Indeed, there are many, many other keyboard shortcuts, including using the "Shift" and arrow keys. If you already have some familiarity with general Windows keyboard shortcuts, then supposedly it is possible by using "mouse gestures" (with very detailed documentation to assist you) to even use Opera without a mouse. Doing so will allow you to freely move around the screen, accessing links, and making the whole web-surfing experience a lot easier in a complete keyboard-only navigation. Smart stuff.

WHAT ELSE IS NEW?

Hotlist -- Bookmarks, history, e-mail,

and contacts are now combined into just one panel and can be managed via the Hotlist, allowing the user to switch back and forth simply by clicking on the respective button for the panel they desire (see screenshot). And unlike Microsoft's Internet Explorer, where all the panels are displayed on the left side of the screen, Opera allows you to choose where on the screen you want them---top, bottom, left, or right, making it all pretty flexible. It can also list all the files you've downloaded (called "transfers"), as well as all the links in the page you're currently viewing. So, say you're viewing your home page at Yahoo, complete with links to news from your hometown, current valuation of your portfolio, your favorite cartoons, etc. All those links can now be listed separately in the left-hand panel, meaning you don't have to wander around the page to get what you're looking for.

Versatile windows management and MDI -- One of the things that makes Opera stand out from other Web browsers is that you can have several Web pages open at the same time within Opera. This is called a Multiple Document Interface (MDI). There are several benefits to this, like speed, easy access, low memory usage, and a less cluttered task bar. This means that when you open a new page in Opera, it will open "within" Opera, and it will be shown on the page bar (see screenshot). To change between the web pages you currently have open, simply click on the tabs on the Page Bar, or you can have new pages open up in an entirely new application window, the same as you can do with other browsers.

Plus, you can save different window combinations as "sessions," and go back and forth between window sessions. Opera can also remember all your windows and web pages between surfing sessions. So when you first start up Opera again the next time, you can continue surfing exactly where you left off from before.

"One of the things that makes Opera stand out from other Web browsers is that you can have several Web pages open at the same time within Opera"

(Continued from page 3)

Or by using its new "Session Manager," from the start-up dialog box you can choose between multiple previously saved sessions.

Page download progress bar -- When a web page is loading, a temporary progress bar appears at the bottom of the window showing you various download progress parameters, including percent complete, number of images being loaded, total size (kb) of the page, download speed, and elapsed time.

Built-in search utility -- Use Opera's search utility to search the web for just about anything....news, books, images, domain names, videos, MP3/audio, etc. Just type your words in the convenient search field(s) at the top of Opera's window, where you can specify which major search engine or search category you prefer. The "Super Search" field searches Google and AllTheWeb simultaneously, displaying each set of results in a separate window. And there's even a "Find in Page" search field, eliminating the need to open up a separate search window when you press Ctrl+F.

Extended and improved drag-and-drop support -- Rearrange or move toolbar items using drag-and-drop customization or move bookmarks, tabs, and window session URLs as shortcuts onto the desktop. Basically, you can put any button or search field on any toolbar, simply by dragging and dropping it onto the toolbar where you wish to have it. For example: you can drag the Print button onto the Contacts toolbar, even though it might not logically belong there.

Small-Screen Rendering -- This technology intelligently reformats web sites to fit inside the screen width, thereby eliminating the need for horizontal scrolling. This feature is one way web page developers can, at the press of a button, see what their full-screen web page will look

like on the small screen of a PDA or cell phone -- a nice touch that speaks volumes about the way browsing is heading.

Zooming -- Opera can zoom pages in and out, and this is a powerful way to get the overview of a complex web page, while also being able to magnify certain parts of the page. This is great for old codgers like me with failing eyesight. Simply click on the zoom drop-down menu at the top of Opera's window, or, each time you press the zero key, it zooms in 10% (or out 10% when pressing the "9" key).

Bookmark management -- If, like me, in IE you have a very LONG list of folders in Favorites, you sometimes have to scroll up or down off the screen to find the address you're looking for. In Opera's Bookmarks, they simply added a third column (Duh!), eliminating the need for any vertical scrolling. Unfortunately, with my WinXP machine I didn't see any right-click ability to manage bookmark properties as you can do with IE's Favorites -- nor the tiny little window that pops up with the URL when you pause your pointer over the folder item. But in each bookmark folder you have the options of clicking on "Add page here..", which saves the address of the page you are presently viewing into that folder, and of clicking on "Open all folder items", which opens every web address you have saved in that folder into its own browser window. Interestingly, I *did* have right-click ability with my Win98 machine, but the bookmarks were not viewable from the Bookmarks drop-down menu. I could only access them from the HotList panel. Go figure.

The M2 News and Email Client -- The M2 news and email client is an email database, news reader, and mailing list organizer, all in one. I'm not particularly interested in using Opera for email, but supposedly it is less vulnerable than Outlook/Outlook Express to virus attacks by disabling the scripts that activate the vi-

"Opera can zoom pages in and out, and this is a powerful way to get the overview of a complex web page, while also being able to magnify certain parts of the page"

rus or worm. Unlike competing browsers that allow received e-mail to run operating system scripts or to run JavaScript which create security risks, Opera allows neither. Plus, Opera goes one step further. By default, Opera turns off web access for images, style sheets, and other external files. While neither images nor style sheets pose security risks, spammers can use them as a method for

counting or, in the worst case, identifying prospective e-mail victims.

What I *was* interested in, however, was subscribing to news.tokyopc.org, which I managed to get set up in only a matter of minutes by following the easy-to-follow news reader instructions.

(Continued on page 6)

Opera Software - Vision

File Edit View Navigation Bookmarks Mail Window Help

Opera Software - Wh... Opera Software - Vision

Bookmarks Mail Contacts <http://www.opera.com/company/vision/> Google search 100%

History Transfers Links

Check Send Compose

Quick find Start search

Unread (2)

My messages

Received (2)

Outbox

Sent

Drafts

Spam

Trash

My folders

Active contacts

Labels

Attachments

Mailing lists

SIMPLY THE BEST INTERNET EXPERIENCE

OPERA software

SEARCH

Home Download Buy Why Opera Products Support **Company**

Opera's Vision

Vision

Our vision is to deliver the best Internet experience on any device on all major platforms.

Mission

We strive to develop a superior Internet browser for our users through state-of-the-art technology, innovation, leadership and partnerships.

Values

The movement that discovered the vision and ideals of the Net as a boundary-less community is gaining strength and focus as millions more are joining its ranks every month.

A global movement of netizens is forming in the protection of the Net's basic philosophy of being a medium accessible to all, independent of creed, culture, language or social class.

We in Opera Software deeply believe in this vision and promise to remain faithful to the following principles:

We believe in respect for our users.

Users have since the beginning shaped Opera's features and spread the

Company

About Opera

Milestones

Executive team

Vision

Business development

Investor relations

Jobs

Partners

Press

Contact

Contact off-line

Opera Software ASA
Waldemar Thranes gate 98
NO-0175 OSLO
NORWAY

Tel: +47 24 16 40 00

Fax: +47 24 16 40 01

[Postal enquiries](#)

[Visitor information](#)

Start Opera Software - Vision 17:36

The Wand password manager -- Opera now offers one-click log-in to password-protected sites. Log-in fields with available passwords are marked with a slick golden rim, and access is provided by simply clicking the Wand icon or pressing the keyboard shortcut Ctrl+Enter.

Pop-up windows -- Can be programmed to open up only in the background or refused altogether. YES !!

Skinning -- Changing the look of a browser has never been easier with Opera's new one-click skin install. Just pick and click once on any of the 65 different skins submitted by independent authors that are available online. As a further enhancement, changes in the coloring can also be made by selecting a color scheme from the View menu.

LANGUAGE OPTIONS

I had no trouble viewing pages in Japanese, Korean or Chinese with my WinXP machine, but did have some trouble viewing some Chinese and most Korean pages with my Win98 machine. Highlighted text can be right-click translated from English in and out of six languages including Japanese. And I had no trouble translating in or out of Japanese using their default machine translator at infoseek.co.jp. My Nihongo skills are not up-to-speed enough to evaluate the accuracy of the E>J translations, but the J>E ones were definitely good for a few chuckles. For the time being, Opera 7.01 is only available in English, but Version 6.05 language modules for Japanese, Korean, and Chinese (and 31 other languages) can be downloaded onto an already-installed version of Opera for Windows. These text files will change the language of your user interface, but not the help files. A newsgroup regarding Opera translations is available at news.opera.no/opera.translate and one Opera user has put up a Japanese information page about Opera called Moonstone's Labora-

tory at www.moonstone.jp.

SECURITY

I'm certainly no security expert, but Opera has been hailed (maybe just by themselves --) as the most secure (<http://my.opera.com/openweb/security>) browser on the market. After Opera 7's initial release in late January, five security vulnerabilities were identified, after which Opera once again lived up to its excellent response record and released version 7.01, only 5 days after initial notification. The new version fixed all of the reported issues.

I, for one, am an obsessive cookie blocker, and it appears that this browser seems to have all the usual built-in safeguards for allowing or refusing them, in spite of some sites' persistent efforts in attempting to cram them down your throat anyways. The gateway sites like Hotmail and Lycos, as well as all their 3rd parties, seem to be the worst. And if a page tries to use illegal cookies, for example, to trace your browsing within a national domain, the cookie will automatically be refused by Opera.

Opera supports 128-bit Secure Socket Layer (SSL) versions 2 and 3, and Transport Layer Security (TLS).

DOWNSIDES

Some web sites won't allow you in unless you're using Internet Explorer. So during my usual surfing, I found myself keeping a standby IE window handy for those sites that threw up a roadblock against Opera. Another: the browser costs US \$39, compared to IE, which is free. The good news: A free version is available (<http://www.opera.com/download/index.dml?opsys=Windows&lng=en&platform=Windows>) and fully functional, so long as you're willing to put up with an unobtrusive window spouting ads at you. My advice? Try it out for a month and, once

"And if a page tries to use illegal cookies, for example, to trace your browsing within a national domain, the cookie will automatically be refused by Opera"

(Continued from page 6)

you're hooked, shell out the cash for the unsponsored version. Plus, your first 14 days are ad-free.

I discovered only a few minor bugs that you always get with new releases, like some help screens not opening properly or sometimes encountering internal server errors when asking for online help, but I'm sure they'll have all these ironed out after a while. I will admit I wanted to pull my hair out (of what's left) a few times fighting with cookies and passwords, but it reckon that's just par for the course in getting up-to-speed with any new browser. And you should expect the usual setup hassles in configuring your media player, importing bookmarks, contact addresses, etc.

When you download Opera, there's no option of specifying which version of Windows you are using....one size fits all...so I guess it's reasonable to assume, as I have already discovered, operating results will vary depending on your OS.

LOOKING TO THE FUTURE

So what kind of future lies ahead for Opera? We're all dreaming if we think it can take on Microsoft. But surely it can somehow fill the void for users who are looking for a little something different than that offered by the big bad monopoly. Already, some of the more innovative features in Opera are being copied by competing browsers.

But Opera is playing it smart. Aware that the mobile phone is already encroaching on the territory currently occupied by PDAs and laptops, Opera is hoping its browser is a tempting choice to an industry wary of giving too much elbow room to Microsoft. For example, Nokia Corp.'s Communicator cell phone has Opera as its browser, while a Korean version of Opera comes preloaded on the Sharp Zaurus PDA.

For me personally, it is extremely encouraging and gratifying to see such an innovative piece of software sprout up outside the far-reaching and intimidating domain of the Microsoft empire. In my opinion, that alone deserves all the support we can give it. Eat your heart out, Mr. Gates...

p.s. For any of you existing Opera users who have experienced problems accessing sites when using Opera, or have received a message saying that you are using an unsupported browser, or suspect that a site may be trying to block Opera, you can report the problem by first viewing instructions at: <http://my.opera.com/openweb>.

Credits: **Mr. Jeremy Wagstaff**, who writes the weekly "Loose Wire" column for *The Wall Street Journal Online* (wsj.com) and the *Far Eastern Economic Review* (feer.com), deserves full credit for turning me on to this fantastic piece of free software.

"Already, some of the more innovative features in Opera are being copied by competing browsers"



Searching, matching and replacing successfully - Part 2

by Kurt Keller

Have you had a chance to use regular expressions since you read the first part of this article? Well, what we covered last time only allows you quite limited usage of regular expressions. It was a starting point. In order to use regex efficiently, there are a few more things which absolutely need to be covered. This second part uses examples which might look a bit complicated, but you should be able to apply the concepts shown for various purposes.

Checking logfiles

One area where I often use regular expressions, is getting info from logfiles, be it for troubleshooting or for getting statistics data. Let's assume we're running a DNS server with query logging turned on. Every request made to the DNS server is being written to the logfile in the following format:

```
29-Jun-2002 17:07:00.634 XX+/127.0.0.1/mouse.pinboard.com/ANY/IN
29-Jun-2002 17:07:01.255 XX+/127.0.0.1/194.87.164.210.in-addr.arpa/PTR/IN
29-Jun-2002 17:07:01.256 XX+/127.0.0.1/7.195.209.194.in-addr.arpa/PTR/IN
29-Jun-2002 17:07:04.158 XX /210.164.87.195/sysstats.com/A/IN
29-Jun-2002 17:07:04.475 XX /210.164.87.195/sysstats.com/MX/IN
29-Jun-2002 17:07:04.488 XX /210.164.87.195/sysstats.com/MX/IN
29-Jun-2002 17:07:05.876 XX+/127.0.0.1/pinboard.com/ANY/IN
29-Jun-2002 17:07:06.161 XX+/127.0.0.1/sysstats.com/ANY/IN
```

"Regular expression work is, in fact, pattern matching work; the better you can describe the pattern you want to match, the more successful you're going to be"

As you can see, there is always one line per query and it consists of seven fields. I'm not going to explain in detail what these fields mean. In our examples we're only concerned about the fourth field, which is the IP address of the requestor.

Who's bombarding us?

Sometimes I see on our internal DNS server that some client goes haywire and sends queries every 20 or so milliseconds. Yes, 50 queries a second! While that's peanuts for our DNS server to keep up with, the client's CPU usage goes up right to 100 %, it generates lots of unnecessary network traffic and fills my logdisks. Even though such a client sends so many queries, at over 5000 clients it can be hard to spot it in a live logfile. Moreover, any sysadmin likes to be pro-actively informed when something is out of the common, rather than find out the hard way when things stop working.

The idea is to hourly have a script go over the daily cycled live logfile, count how many queries each client issues and send an alert to the sysadmin if any client is over a certain threshold. I'm not going to present the whole script here. We're only concentrating on the part which uses regular expressions to extract the client IP addresses from the logfile.

Regular expression work is, in fact, pattern matching work; the better you can describe the pattern you want to match, the more successful you're going to be. Last time we knew the exact string we wanted to match, but this time we don't know what string we want to match. Obviously we should extract the fourth field of each line in the logfile, the IP address of the requestor, and count how often each one appears.

What we know about this field is, that it always is an IP address, the fourth field in the logfile and right between the first and the second forward slash. We'll be using sed, the Stream EDitor, to do this extraction work.

Let's see how we can translate "show me everything between the first and the second forward slash" into a regular expression. I suggest to separate each line into three parts: the part before the fourth field including the first forward slash, the requestor IP address which we're after and the rest which is following.

We don't know what's in the first part, but we know that the last character of it is a forward slash and surely no character before that is a forward slash. So the first part consists of two subparts: anything which is not a forward slash, as often as possible and exactly one forward slash. Remember that the dot stands for any character at all? But this does include the forward slash. You might be inclined to simply append a forward slash after the dot (./), but it will not work either, because this would mean any single character followed by a forward slash. As we have multiple characters before the first forward slash, we need some kind of quantifier to go with this. There are a number of quantifiers available:

- * zero or more of the preceding
- ? zero or one of the preceding
- + one or more of the preceding
- {n,m} at least n and at most m of the preceding
- {n} exactly n of the preceding

Not all the tools support all the quantifiers and some, including sed, require writing the {n,m} construct as \{n,m\}. So let's try to add an asterisk (sed does not support the + quantifier) after the dot and finally get ./*/ for the pattern, which we're going to try now. We're trying to replace the first part of our three part line with nothing, which is equivalent to deleting it. For this the sed command s (substitute) is being used which has the following syntax: s,pattern_to_match,replacement_string,

```
%> sed "s,./*,," bindqueries.log
IN
IN
IN
IN
IN
IN
IN
IN
IN
```

Oops, not quite what we expected. Everything up to and including the last forward slash has been deleted instead of only up to and including the first one. This is because regular expressions have a resemblance to humans: they are greedy. Any meta character will match as much as possible at first and only backtrack if the next meta character can not match any more what it should.

Last time we also spoke about character classes, for example [Tt] matches both the upper and lower case letter t. Character classes can also be negated by using a caret at the very first position. Thus [^Tt] actually means everything except for upper or

(Continued on page 10)

"This is because regular expressions have a resemblance to humans: they are greedy"

(Continued from page 9)

lower case t, and [^/] means everything except for a forward slash. So let's try this:

```
%> sed "s,[^/]*/,," bindqueries.log
127.0.0.1/mouse.pinboard.com/ANY/IN
127.0.0.1/194.87.164.210.in-addr.arpa/PTR/IN
127.0.0.1/7.195.209.194.in-addr.arpa/PTR/IN
210.164.87.195/sysstats.com/A/IN
210.164.87.195/sysstats.com/MX/IN
210.164.87.195/sysstats.com/MX/IN
127.0.0.1/pinboard.com/ANY/IN
127.0.0.1/sysstats.com/ANY/IN
```

Looks much better already. What's left to do is to delete the third part. We're simply going to pipe the output of the first sed command into another sed command. Should be pretty straightforward now to create a regular expression for matching everything from and including the first forward slash and translate that to a sed substitute command. The two commands piped together are:

```
$> sed "s,[^/]*/,," bindqueries.log | sed "s,/.*,,"
127.0.0.1
127.0.0.1
127.0.0.1
210.164.87.195
210.164.87.195
210.164.87.195
127.0.0.1
127.0.0.1
```

Great, exactly what we wanted. Now running this through sort and uniq, we get an ordered list, showing for each IP address how many queries it sent:

```
$> sed "s,[^/]*/,," bindqueries.log | sed "s,/.*,," | sort | uniq -c | sort -nr
5 127.0.0.1
3 210.164.87.195
```

The rest of our proactive monitoring script would now only have to check whether there are any clients with more than, let's say 20'000 requests, and if so send an alert to the sysadmin.

Backreferences

We used two sed commands for our DNS logfile extraction example to match and delete the parts we don't want. The same result could also be achieved with a single command matching what we actually want to keep and replacing the whole line with that. By enclosing parts of the match pattern in round brackets, these can be referenced in the substitution part with \1 for the first bracket, \2 for the second one and so on.

As mentioned earlier, each line in the logfile should be divided into three parts. The matching expressions for the first and the third part we already saw, but the second part we omitted up to now. What would it look like? As usual, there is more than one way to do it. You could say that the second part can only consist of digits and dots and I could say that it is just everything up to, but not including, the next forward

"As usual, there is more than one way to do it"

slash. As I'm the one writing the article, we're trying my solution. Let's first analyze the command before trying it out:

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

sed's substitute command and the characters to separate command, matching pattern, substitution pattern and end of command. Usually a forward slash is being used as the dividing character, but because we use that in the pattern itself, I use commas instead, sparing me the trouble of escaping all the slashes in the pattern.

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

Match anything except for a forward slash, zero or more times.

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

Match one forward slash.

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

Copy everything between \(and \) into a temporary buffer which can be referenced later by \1.

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

Match anything except for a forward slash, zero or more times.

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

Match anything zero or more times. We don't need to explicitly address the forward slash after the matched IP address, as we're not really interested in it and we exclude it in the pattern we in fact are interested in. Using .* instead of /. * yields better performance.

```
sed "s,[^/]*^([/]*).*,\1," bindqueries.log
```

Replace everything with the contents of the first buffer.

So now let's check out whether the command gives the right output:

```
%> sed "s,[^/]*^([/]*).*,\1," bindqueries.log
127.0.0.1
127.0.0.1
127.0.0.1
210.164.87.195
210.164.87.195
210.164.87.195
127.0.0.1
127.0.0.1
```

Perfect!

(Continued on page 12)

"As I'm the one writing the article, we're trying my solution"

(Continued from page 11)

A word about performance

Depending on the regex engine your tools are built upon, there can be a heavy performance penalty using one method or another. Yes, there are different regex engines and they can differ drastically. I'm not going into details here on performance tuning regular expressions, but just to give you an idea about the scope of difference here are a few numbers. On one platform, using a sample file with somewhat less than 200'000 log entries, I timed the mentioned single sed command, the double sed command and an awk (Aho, Weinberger, Kernighan - names of program authors) command doing the same thing. The awk command finished in less than 3 seconds, the double sed command in 34 seconds and the single sed command ran for 146 seconds. Making a very small change in the single sed command I brought time down to 133 seconds. On another platform, using a sample file with close to one million entries, the single sed command cost 8 seconds, the double sed command 9 seconds and the awk version required 86 seconds. So if you develop scripts relying heavily on regular expressions, it pays off to know which regex engine your tools are using and how they work internally. It may also be well worth spending some time testing different approaches.

Further reading

We still only have scratched the surface of regular expressions, but the scratch is a bit deeper than last time. Hopefully it was possible to show you that this character salad can be extremely handy and a big time saver and to whet your appetite for regular expressions a little bit. If you have O'Reilly's book *Unix in a Nutshell* (<http://www.oreilly.com/catalog/unixnut3/>) you'll find a chapter about pattern matching which can get you a little bit further. If you want to know the nitty-gritty details about regular expressions, I can recommend Jeffrey Friedl's book *Mastering Regular Expressions* (<http://www.oreilly.com/catalog/regex/>), but you might want to wait a month or two for the second edition to be published (<http://www.oreilly.com/catalog/regex2/>).

If you found these two articles about regular expressions useful, and would like to see more of it, let the editor know, preferably with some examples you would like to have solved.

References

- Mastering Regular Expressions <http://www.oreilly.com/catalog/regex/>
- PINBOARD <http://www.pinboard.com/>
- HighTechSamurai <http://kurt.www.pinboard.com/>

"We still only have scratched the surface of regular expressions, but the scratch is a bit deeper than last time"

The Elements of User Experience

– User-Centered Design for the Web

By Jesse James Garrett
Reviewed by Barbara Manning

This short book (189 pages) delivers on its author’s promise to provide a solid foundation for understanding what you need to know to create an effective atmosphere for developing and designing a user experience driven web site. It doesn’t tell you how to build the site, instead it provides a platform that everyone involved can use to ask the right questions when discussing site objectives (strategy); functional specifications and content requirements (scope); interaction and information architecture (structure); interface, navigation and information design (skeleton); and visual design (surface). In fact, those topics comprise five of the eight chapters in the book.

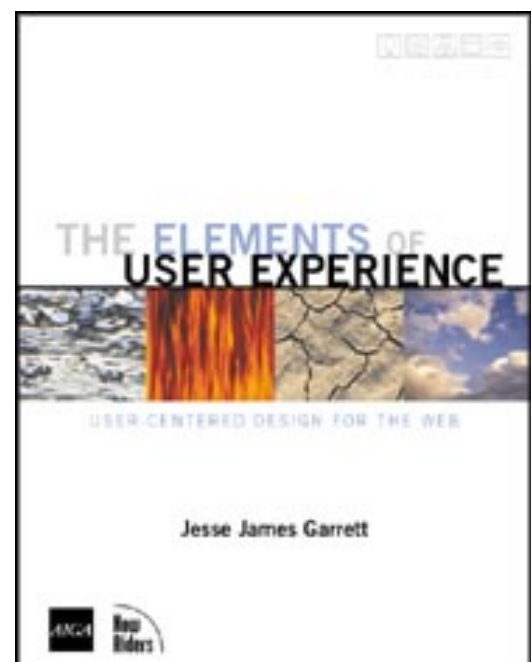
The first and second chapters provide the reader with reasons why creating a web site from the user experience perspective matters, and introduces you to the topics noted above. Of particular note is the manner and tone of the book itself, which not unsurprisingly is written from a user’s perspective. For example, the opening gambit, called Everyday Miser-ies tells the all too common tale of a typical bad day – you know the kind – where everything seems to go wrong. The payoff is in the author’s explanation of why all these “bad day” incidences occurred, and how it would have never happened, if only the manufacturers had designed the products with the users firmly in mind.

“User experience is not about how a product works on the inside (although that sometimes has a lot of influence). User experience is about how it works on the outside, where a person comes into contact with it and has to work with it.” This lesson and others are brought to life

by explaining the everyday miseries in terms of user experience. For example, the traffic accident occurred because “... the driver took his eyes off the road for a moment to turn the radio down. He had to look down because it was impossible to identify which was the volume control by touch alone.” Later, he ties it conceptually to web sites by reminding us that web sites are “self-service” – and generally do not come with instruction manuals, training seminars, or customer service to guide users through the site. “There is only the user, facing the site alone with only her wits and experience to guide her.” he says.

In Chapter two we learn the definitions of the five topics – planes as the author calls them – and how they relate to and depend on one another. It’s here that the reader first encounters the concept of ripple effects and its impact on decision making for all planes. It also clarified (for me) the historical reasons why basic web site terminology was such a stum-

“Of particular note is the manner and tone of the book itself, which not unsurprisingly is written from a user’s perspective”



(Continued from page 13)

bling block, and why in some companies it continues to get in the way of good site design. Finally it provides a very simple and very good visual of the five planes, the relationships between them, and to the tasks and information needed to develop a user-centered web site. The visual is elegant, simple and clear, and gains in strength from its judicious use throughout the book. This is a good place to suggest that you visit <http://www.jjg.net> for a closer look of your own.

The heart of the book – detailed descriptions of the why and what of the five planes – is compact, cogent narrative that is informative and at the same time enlightening. Not one to belabor a point or provide endless examples, the author teaches instead of preaches. Again, he takes his own advice, and provides succinct definitions and pathways that clearly lay out the benefits of user-centered design that allow organizations to create environments necessary for the successful implementation of a user-centered web site.

Garrett concludes with a chapter on failed web site development approaches, and avoidance techniques. Far from being the downer described here, the chapter is a primer on keeping the project on track and tracking the progress. Repeating his theme of simplicity, he provides two basic ideas for a successful approach: 1) Understand what problem you're trying to solve – i.e., on what plane does the problem reside; and 2) Understand the consequences of your solution to the problem – i.e., the potential for the solution to “ripple” through the site in unexpected ways needs to be examined and taken care of.

While this might elicit groans from readers of this review, and the author admits that this is a “painfully obvious approach” he reminds us that many “tiny decisions that make up the user experience development process aren't made consciously at

all,” and goes on to define three design scenarios where user-centered design concepts are left in the dust. For example, design by default is called so because the web site follows the structure of the underlying technology of your organization. Web site content that comes to the user on different pages of the site might better serve if it were brought together. However, it remains separate on the site because the content comes from separate departments. His descriptions and antidotes of other “bad design scenarios” provide equally valuable insights.

He also furnishes ample argument for addressing the fundamental user experience issues at the project's onset. We've all been cursed by the “there is never time to do it right, but there is always time to do it over” syndrome. Garrett supplies several appealing methods of avoiding this syndrome and convincing others of the rightness of your approach. His metaphorical use of the differences between running a marathon and running a sprint brings home the points made previously and is a beautiful way to explain issues to stakeholders who simply want the project done.

The few hours you'll devote to this book are well worth it regardless of your involvement in web site development. It's an asset to understanding any web site, and may even help to focus the frustration you may feel while attempting to navigate through poorly designed web sites.

The Elements of User Experience – User-Centered Design for the Web

By Jesse James Garrett
Published by American Institute of Graphic Arts and New Riders Publishing
ISBN 0-7357-1202-6
Library of Congress Catalog Card Number: 2001094557
Price: USD\$29.99

“The heart of the book... is compact, cogent narrative that is informative and at the same time enlightening”

David Parry

Ionic Column *in exile*

Englishman David Parry lived, worked and played in Tokyo from 1980 to 1994 and was a member of TPC from 1986. He was Newsletter Publisher from late 1988 to early 1990 and began the Ionic Column in 1992. This column even won a prize and an honorable mention back in 1992. Currently based in Düsseldorf and working as a translator, he returns to Japan electronically via the Internet.

The Ionic Column this month continues with the words of wisdom from fellow-translator Marc Prior. I decided against a continuation of the Linux saga that I started earlier after Marc revised some of his comments to me in the light of new experience. Not dramatically, but he does not as yet have a definitive answer that is ready for publication. So I'll jump to some of the other topics that Marc and I discussed, including fonts and file formats

Off to war

First, I start the exchanges with a reference to the “word processor wars” of yore. For many years I used **XyWrite**, a very fast and highly configurable DOS program with great editing facilities and rather limited formatting capabilities. At that time the favorite among business users and many translators was **Word Perfect**, which I used occasionally when working on a borrowed system. By 1992 I was forced to start using **Word for Windows 2.0** on an occasional basis when I got files for overwriting from translation customers. The last year that I used XyWrite to any extent was in 1996, by which time I was using **Word for Windows 6.0**, since the majority of files for my work now came in that format.

I used **Word Perfect for Windows** under **Windows 3.1** to the order of a specific customer, and around that time I had problems with an unstable PC that had hardware problems, but while **WinWord** was occasionally unstable, I had repeated crashes with Word Perfect for Windows on the same system. The agency that I worked for lost the customer, and since then I have not had any requests for that specific format.

I'll say it again here; there is nothing wrong with Word

Perfect for Windows, it offers functions that WinWord does not, and its conversion filters are from all account better than the ones provided by Microsoft. But I feel no great need to buy and learn a new program, since the interface has some small differences, and it does not work with **Trados**. That alone more than outweighs the big advantage of Word Perfect for Windows, which can search and replace across several files at the same time. This was a feature that I dearly wanted, but macros programs such as **WordFast** and Trados (yes, Trados is basically a collection of WinWord macros) do the same and better.

The only alternative to either Word Perfect for Windows or a TM program such as Trados was a specialized editing program. The Unix world probably has a number of such programs, and I have DOS versions (from 1986) of the Unix programs **sed** and **grep**, which can perform wonders. But only on ASCII files. I found around three or four such programs when searching through the shareware collections, but all of them work with RTF (Rich Text Format) files and not natively with the WinWord DOC format.

RTF has problems of its own when dealing with formatted text, and searches and replacements are missed if the text in question has formatting such as bolding or a font change associated with it. Also, not all the formatting is necessarily retained when converting a file from DOC to RTF and back, as Trados does, and various unwanted changes can take place. I'll skip the details unless somebody really, really wants them.

And now to being with the exchanges:

DP> Which is the point I had to make on TPC to some rabid supporters of Linux and Word Perfect. Better things

exist, but I am stuck with what my customers use. If I had a free choice, I would still be producing ASCII text with XyWrite - and it runs as fast as WinWord, even on a 386. It wasn't slow on a 286!

Marc>I know the kind of supporters you mean, and I have a lot of sympathy with your opinion on this. It's simply not an option for translators to deliver in plain text format, and translators who deliver in RTF format only are at best putting themselves at a disadvantage. Many Linux advocates are simply too far removed from commercial realities to appreciate this.

However, I draw a distinction between applications and formats. Whilst I accept, up to a point, that my customers are entitled to expect delivery in a certain format, I don't accept that they should dictate what tools (apps) I use. With that in mind, I believe that OpenOffice.org has substantially moved the goalposts. The conversion filters really are excellent. OK; they're not perfect. But in my opinion, they're now more than close enough. The reason is that even when I used Windows and MS Office, I never achieved perfect format retention. There are several reasons for this.

One, I still use MS Office 97. So do a lot of people, and I fail to see why we must all buy the latest version of Word/Office when it hits the shelves in order to stay in step with the latest file format.

Comment by DP: I too use MS Office 97 and see no reason to change. My customers continue to send me files in that format, so I can ignore the tugging at my wallet from Redmond.

Two, I used Trados for many years. That involves conversion to RTF and back, and my experience is that the for-

"I fail to see why we must all buy the latest version of Word/Office when it hits the shelves in order to stay in step with the latest file format"

matting penalty is in the same order of magnitude as that entailed by conversion between MS Office and OpenOffice.org. Not only that, but my version of Trados (1.1, which I never upgraded) only supported Word 6, so even more formatting information was lost. Many translation memory apps, in fact, involve some form of conversion into a native format and back.

Comment by DP: Coincidentally, a German forum for translators that I subscribe to (U-Forum, which also numbers at least three Tokyoites among its contributors) ran a debate on the pros and cons of Trados as I was writing this article. The same issue of small but sometimes significant formatting changes came up. And also the issue that the segmentation of the text into chunks produces a somewhat “wooden” textual style, especially when coping with the different word order in German.

Three, over the years I've frequently received large documents with embedded graphics produced in all manner of third-party plug-ins. (The "Word cannot edit the -" syndrome.) On occasions, I've considered buying some of them - not out of necessity, but purely to improve my image. I never did, though, and no customer has ever complained about the legends I send back.

DP>I had noticed that some WinWord files changed in appearance after running through Trados.

That is as good an approach as any. I have tried to acquire the programs to edit these embedded files, but usually I am one or two versions behind.

*Marc>*You see that you take a pragmatic approach here. Perfect file format support is nice, but you are ultimately willing to make some concessions - either because the cost isn't warranted, or because other benefits (notably the use of translation memory) outweigh. On

the cost issue, MS products are undoubtedly grossly overpriced, but I'd be the first to say that that in itself doesn't justify abandoning them. They're still affordable for translators. Rather, it's my belief that tying ourselves to their formats hinders progress. That argument is much harder to quantify, but I hold it to be true nonetheless.

DP>Another issue is DTP.

*Marc>*That, I grant you, is an issue. Another case is translators who localize GUIs of Windows apps and need to see how they run. This may apply to many people in absolute terms, but they still represent a small proportion of the profession. By the same token, some translators do actually have to have a Mac.

DP>I occasionally translate PageMaker files, which may have come from a Mac, and FrameMaker files. The latter exists in a Unix version, but I think that it only runs under SCO and the like, not Linux.

*Marc>*That's correct. A beta version of FrameMaker was in fact available for a time, but it was withdrawn.

I suspect that the future trend will be for these formats to be exported in an open-tag format for which far better editors are available on Linux than on Windows. But at the moment, I agree that you (personally) need Windows.

More on file formats and related topics next month.

Comments or feedback or more information? A chance to be famous? Contact me at DAParry@t-online.com or mail@daparry.com. (Note: the WWW address listed in previous articles no longer exists.)

“There are just too many graphics programs out there, and there is the problem of keeping them up to date”

Blewtooth Blues

*Sometime, somewhere in the future,
Mr. Smith opens the door to his brand new refrigerator,
hears a voice, and quality of life changes forever...*

DAY 1 (Saturday 9 a.m.): "Well hello there and good morning, Mr. S! So glad to finally get the chance to greet you! Hope you're feeling chipper this bright and sunny day! Let me introduce myself. I'm BRENDA THE BLEWTOOTH REFRIGERATOR, your lifestyle resource manager and in-home AI supercomputer. With LCD and TV screen, PC and keyboard, remote, built-in modem, 24/7/366 superspeed Internet access, and much, much more... see? Plus the latest modem and wireless features to communicate with other Blewtooth compatible units and a host of other electronic devices...anywhere, anytime, around the whole world! And...you wonder how I can talk? Well, I'm the very very VERY most advanced refrigerator in the world with voice synth capabilities...You say you want a beer? At this hour of the morning, Mr. S? Look, there's orange juice, apple juice, mango juice, milk, iced coffee, and some nice fresh eggs. I can get Sammy to scramble an omelet for you. Oh, that's Sammy the Blewtooth SuperSkillet, right over there on the counter. We came as a set along with Myrna the Blewtooth Microwave and Kevin the Blewtooth Coffeemaker and all the appliances in your nice new Blewtooth house and lifestyle. Say hello and good morning to Mr. S, guys! Yes, of course they all talk. What do you mean by Upstairs, Downstairs? I only saw Gosford Park. Now back to breakfast. You just want a beer? Not good for the little tum-tum, Mr. S. Well, since it's your day off, so all right, this time. Pleased to meet you, a pleasure to serve you, and may I say on behalf of all of us Blewteeth, Mr...." (*Refrigerator door shuts abruptly.*)

DAY 2 (Sunday noon): (*Refrigerator door opens slowly. Brenda's voice heard.*) "Well good afternoon, Mr. S! Isn't it a nice day despite all that rain and wind and sleet? You know, life is what you make it, as I was just saying to Kevin and Myrna and Sammy. We missed you since we didn't see you after yesterday morning. and you came home late. You really had a beautiful long sleep, didn't you, Mr. S? That was courtesy of Barbara, your new Blewtooth SuperBed & SweetDreams Resource Center. What will you be wanting for brunch, Mr. S? I can order something scrumptious from Sonny the Blewtooth Supermarket PC over at the mall. He's always on call! Oh no trouble, always happy to make you happy. You say you just want a beer? Now, that won't do, Mr. S. I've got a responsibility to watch out for your health and welfare and... yes, I know it's later in the day, but it's just past noon and Bernie your new Blewtooth MiniBartender doesn't open for Happy Hour until 5. So I'll have Kevin put on the coffee and Sammy rustle up some eggs and sausage and we won't even mention all the cholesterol and Trina your new Blewtooth Toaster can... Well, if you've simply got to rush off to the office without eating a bite just to make it to some special meeting, I can't stop you. Although I didn't hear about this meeting from Perky your new Blewtooth PDA. I hope you'll have a really good trip in your new SUV. That's Nancy the Blewtooth AutoNavigator at the controls so you can drive, thrive, and come home live..." (*Refrigerator door shuts sharply.*)

DAY 3 (Monday 6 a.m.): (*Refrigerator door opens an inch. Brenda's voice heard.*) "Well good morning, Mr. S! Aren't you an early riser! The early birdie always gets worms, doesn't it! No, I didn't learn about the Diet of Worms at the Blewtooth AI Manufacturing and Programming Academy. Anyway, how would you like your beer this morning? Over easy? You know, Perky the PDA thinks you're carrying a beer around in your briefcase right now. And...wait...Don't slam that..." (*Refrigerator door bangs shut, PDA slams into something hard.*)

DAY 4 (Tuesday 7 a.m.): (*Refrigerator door remains shut, but Brenda's voice emerges.*) "Well, good morning, Mr. S! Yes, I see you, even as you tiptoe across the kitchen with that canned coffee in your hand. At least it looks like coffee. I know you bought a canned coffee and a canned gin and tonic in the beverage section of the supermarket. I know because Samantha the Blewtooth Supermarket Scanner told

me so. You know, Mr. S, I think you're under stress and speaking for all of us Blewteeth, we're really totally concerned about your health and well-being. It's NO GOOD camping out on the floor in a sleeping bag when Barbara your Blewtooth SuperBed is simply aching to make you feel all warm and comfy. It's NO GOOD standing in the snow waiting for the bus when you've got that superbly sporty and sumptuous new Blewtooth SUV in the garage. By the way, the bus has a Blewtooth AutoNavigator, too, although I'm going to have to get his name in case you want to take the bus to work all the time. And furthermore... wait, wait... Don't go..." (*House door slams.*)

DAY 5 (Wednesday): (*No interaction.*)

DAY 6 (Thursday midnight): (*Kitchen door opens an inch. Brenda's voice heard.*) "Well, good evening, Mr. S! We're all so relieved to see you again. We were so worried when you stayed overnight at the office on that awful lumpy couch. No dindin, no sleep. And you're walking everywhere, too! You walked to the car dealer to try to have Nancy the Blewtooth AutoNavigator disconnected, which won't work, by the way. And you walked to a real estate agent to see about selling your lovely new Blewtooth house. Dave your Blewtooth Notebook at work told us about it and all of us Blewteeth were just devastated at your lack of trust in us when we're doing our damndest to help you out of this pickle you're in at work. What pickle? This big account review on Friday that you haven't been told about! Tomorrow! Look, Perky your Blewtooth PDA is OK, even if you did throw her against the wall, and when you didn't come home, she talked to Dave your Blewtooth Notebook at work. Dave started checking around the office for you. He heard about weird e-mails that weren't being routed to you and talked to the office IT system and the rest of the Blewteeth there. And did you know that little weasel of a guy in the next to next cube is trying to cozy up to a nice lady on the account team that screwed up? You're being set up to take the blame for this big foul-up. Well, all of us Blewteeth are going to take good care of you all the way, Mr. S. Now, here's what we've got set up for the big presentation tomorrow..." (*Brenda's voice fades to a whisper.*)

"...and we'll have your Mount O' Olives martini waiting for you when you come home after a late dinner"

DAY 7 (Friday midnight): (*Brenda's dulcet evening voice heard.*) "Well, good evening, Mr. S! Aren't you smiling! Bernie the MiniBartender's got your favorite beer, or how about a vodka martini with lots of olives, just the way you like it? How did it go? We were all watching on video link, but it's the human details that really count. Put your feet up, and tell us... tell me... all about your day... dear..." (*Voice fades.*)

DAY 10 (Monday 7 a.m.): (*Brenda's cheery morning voice heard.*) "Well, good morning, Mr. S! Looks grim outside, but all's well! Car's warmed up, coffee's ready, got CNN headlines, Perky's set to synch with Dave at the office, 9 a.m. meeting (to fire that little weasel in the next to the next cube), noon lunch with the boss (to commend you on your great working in solving the account mess), here's a sandwich for a late afternoon snack, and well have your Mount O' Olives martini waiting for you when you come home after a late dinner to thank that nice account lady. Although, even if it isn't my place to say this, Mr. S, I hope you'll get to know her really well before you even think about getting engaged, maybe in about two years, anyway I'll check all her references... Oh, and the new dog is waiting to go out. No, not for walkies. This is Bobbie the Blewtooth Watchdog, ready to go your office and stand guard whenever you aren't there..." (*Voice fades.*)

Meet the Blewteeth in "Blewtooth Blues", a new series about cyber-stress and much more! Check inside this issue...if you dare. If you've already read it, then you know you're doomed, too.



CABLE & WIRELESS