Getting the **message**

Nigel Whitfield has relocated and reformed. In this all-new Hands On column, he urges you to crack down on junk email, goes fishing for Perl users, and answers all your net queries.

"Y**ou have been personally selected to receive this** information. Our extensive research suggests it will be of particular interest to you. But don't worry, because in spite of contacting you this time, we promise we won't be doing it ever again. Bear with us this once, however, and your life will be transformed into one of untold riches and splendour."

Internet service providers are becoming more pro-active in fighting the problem of junk mail like this. Sending such mail is prohibited by most online services, and nowadays you are far more likely to find your account terminated for doing it. You're also more likely to find yourself having problems when the internet community decides that the service provider you use is guilty of harbouring spammers.

Earlier this summer, some companies blocked mail to Enterprise customers because of junk mail that had earlier originated from that supplier. Both CIX and AOL now filter out email from known senders of junk messages, as do web-based systems such as BigFoot. And Demon's new web service allows you to kill messages without having to first download them to your system.

There are those, junk mailers among them, who simply see messages as freedom of speech. But they miss the point that we're paying for that freedom: like opening the door, being handed circulars by the postman, and then asked for money to cover excess postal charges.

The internet, although now largely commercial, still relies on co-operation for the smooth transfer of traffic from one part to another: the money you pay for your net subscription is a small contribution to one of the world's largest co-operatives. Junk mail isn't just a nuisance, it's a selfish abuse of that co-operation.

So the next time some junk mail drops into your mailbox, don't reply to it. Don't just delete it: look at the headers and see if you can work out where it came from, and if you can, complain about it. Don't stand for selfishness — and make sure your provider won't, either.

Reading web form input using Perl

Perl is one of the most popular languages for writing scripts to run on your web server; it's relatively easy to learn and most scripts can be moved easily between different systems, including Unix, NT and Windows 95 (Perl for Win32 can be found at www.activeware.com). A good place to start is the book *Learning Perl*, published by O'Reilly. If you can program in other languages, you'll probably be able to pick up a fair bit from the online manuals at
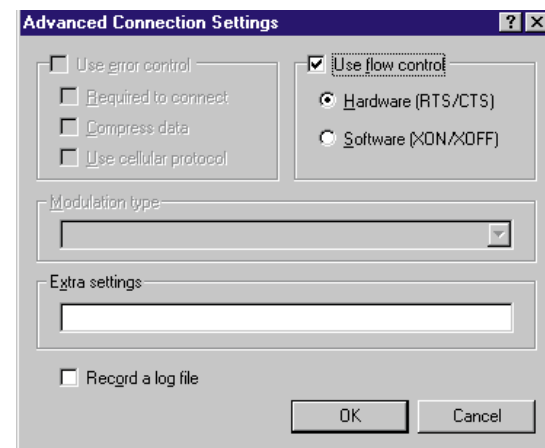
Questions & net.answers

Q **I have a 33.6Kbps modem and I don't know what effect the driver has on its speed. I have experimented with the standard 28.8 driver supplied with Windows 95 and the 14.4 driver that came with the modem, and can't tell the difference. Which driver should I use?**

A Always use the fastest driver you can, although you're likely to see little difference between them if you use basic drivers. The differences come where the driver tells the modem which speed to use and it's unlikely that the Standard drivers included in Windows do that, as it's one area where modems differ tremendously.

The best way to configure a modem is really to do it yourself, or add extra commands to the setup options to ensure the

Fig 1 You can add extra modem options to override the ones that Windows selects in the driver

modem is being set correctly, as you can't see commands that Windows is choosing for you. For extra options, check in your modem manual for the commands to make sure it attempts a connection at the highest speed, select the modem in the control panel, choose Properties, click on Advanced and type the extra commands in the box at the bottom of the window (Fig 1). Just as important is the speed at which the computer talks to the modem. That's the "maximum speed" setting on the control panel, and for a 33.6 modem it should really be set to 115,200bps if you have a fast UART serial-port chip on your system. If you don't have a fast serial port, then the discussion is largely immaterial, as you won't get great performance whichever settings you choose.

Q **I'm writing some web forms and want to use scripts to process the data that's submitted. What's the difference between the POST and GET methods and which one should I use on the pages?**

A The best method to use for almost every situation is the POST method. It's more flexible, you can transfer more data, and, to an extent, it's also more secure.

The difference between the two methods is straightforward. In the GET method, the data from your form is stored in an environment variable called QUERY_STRING. The limiting factor is the size of an environment variable which varies from one system to another, depending on how much memory is allocated to the environment. In the POST method, the same information is sent as input into your script, so there's no limitation to the amount of data; the CONTENT_LENGTH variable tells the script how much there is and exactly that number of characters can be read as input.

For an example of how to decode the information sent to a script, see the section "Reading web form input using Perl" (opposite page). Remember that while the POST method is preferred, it may not be supported by all servers: for instance, the Demon Internet home pages only allow use of GET, which limits the amount of information you can have on a form.

Q **I've enabled ratings in Internet Explorer to prevent my children from seeing some types of site on the internet, but when I try to browse, I find there are lots of sites that I simply can't see. Is there a way around this?**

A Yes, there is a way around the problem: what you've run into is one of the most common problems with the ratings system used by some sites. Since it's a

voluntary system, not all sites include the special tags that are used for rating a page. By default, Explorer won't display pages that don't have ratings, which is fine if you

Fig 2 **You can tell Explorer that you want to see pages with no rating, at your own risk**

want to protect children but awkward for accessing some types of pages, like technical support, which are seldom rated.

To change this option you should choose the ratings options and click on Settings. Choose the General tab and check the box marked "Users can see sites which have no rating" (Fig 2). Alternatively, disable the ratings system for your current setting. Either way, remember that if you don't change the options back, you'll run the risk of people being able to use the browser on your computer to access sites that you'd rather they didn't see.

Q We're getting "Content Advisor" messages when trying to access any web pages via MS Internet Explorer 3.0 (Windows 95). The ISP says this is due to a password having been entered but nobody has knowingly done it so we don't know what it is. They say it is technically possible to edit the registry and reset to "Nil" but cannot provide the way to do it. Trying to access MS help web pages are similarly restricted. Am I doomed to an internet-free life (albeit saving my marriage)? Is

changing the browser the only answer?

A You can fix the problem without having to install a new browser, although a search of the Microsoft web site couldn't come up with the solution to this problem. Trial-and-error has, however, and we've noticed no ill-effects as a result. As it involves changes to the Windows Registry (which is where all the configuration information for your system is stored), you should take a backup, using a tool like the Emergency Recovery Utility.

The program you need to make the changes is called Regedit; you can start by clicking on the Start button, choosing Run and then typing regedit into the box. The Registry Editor is similar to Windows Explorer: to find the information you need to change, open folders in the left hand pane, starting with HKEY_LOCAL_MACHINE, and navigate your way down to

`HKEY_LOCAL_MACHINE\SOFTWARE\Microso`
`ft\Windows\Current`
`Version\Policies\Ratings`

In the right pane, you'll see a display similar to the one in Fig 3: the hexadecimal codes are the password, encrypted so that it can't be broken. You don't need to worry

Fig 3 **Regedit will solve your problem, but incorrect use is risky**

about what it was set to; with the Ratings folder highlighted, press the Delete key and the whole folder will disappear.

Next time you go into Internet Explorer, from Options, choose Security. You'll notice that the system still thinks that ratings are enabled. Click on the Disable Ratings button and you'll be asked to enter a password. You can now use a new password of your choice and you'll then be able to turn the content advisor off and on at will. If you wish to leave it turned off without a password set, simply enter a password, then follow the procedure above to delete the registry entry again.

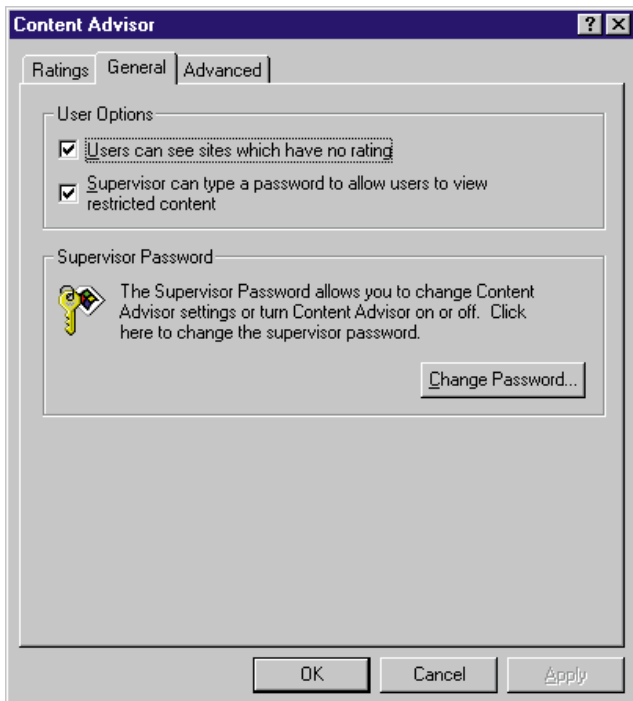Q Is there a way that I can run a mailing list using a standard dialup account from my internet provider?

A In theory, you can run a list. There are versions of list software for machines running Windows and Mac operating systems. You can manage a list by hand, simply by maintaining an alias list in a mail program like Pegasus mail. However, bear in mind that if you want everything automated, you'll need to be able to identify messages for the list automatically, which may not be easy if you only have one mailbox from your provider.

It can be time-consuming to distribute messages to a mailing list, depending on how your provider's email system is set up. It may want to check each address as it's sent out. For some addresses, that could take as long as a minute online. With a large list, that time can soon mount up.

The best solution is to find someone to host a list for you, on a system that's permanently connected to the net. Look for a system with software like Majordomo which can be remotely controlled by sending email from your home account (and via the web) so you can do admin without running up your own phone bill.

If you decide to run a list of your own, invest in an account with a provider that lets you have multiple mailboxes so you can allocate a unique address to the list, making everything easier to manage.
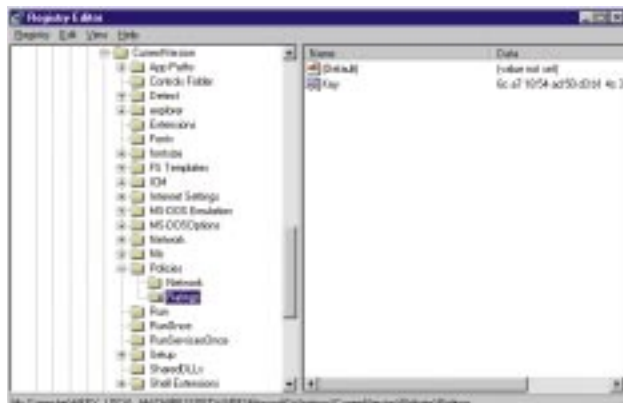
www.perl.com too.

Scripts that create pages on-the-fly are easy: all you need is Print commands to create the HTML you want and you can turn data on disk into nicely formatted pages. But to be really useful, you'll need to read information from the web server and to store it in variables within your script. Some servers provide tools especially for this, but a few lines of Perl can do the trick for every web form.

First, you'll need to use the POST method of submitting data: that means your web page will need a line something like

```
<FORM METHOD='POST' ACTION=
'myscript.cgi'>
```

In your script, you'll need to read the appropriate number of characters from standard input, which effectively means that it behaves as if the form data was typed into the script. The web server uses the environment variable CONTENT_LENGTH to say how many characters there are, and URL encodes the actual data entered into the web page. That means that all the fields from your form are combined into one line, with certain characters represented by a % followed by the hexadecimal ASCII code. For example, a form with a field called "name", into which "Nigel" is entered and a message field with "Hello world", would be encoded as:

```
name=Nigel&message=Hello%20world
```

To make use of the information, you'll need to split the line into parts and change encoded characters back to their original form. The Perl for this is shown above.

Perl does the job

```
1       read(STDIN,$qs,$ENV{'CONTENT_LENGTH'}) ;
2       foreach ( split('&',$qs)) {
3               $_ =~ s/\+/ /g ;
4                       $_ =~ s/%(..)/pack('c',hex($1))/ge ;
5                       ($key, $val) =  split( '=', $_ , 2 ) ;
6                       $i{$key} = $val ;
        }
```

1. This line reads from standard input into the variable $qs, the exact number of characters specified by the CONTENT_LENGTH environment variable.
2. Now we split the query string ($qs) into parts for each field; they're separated by the ampersand (&) character. Since we've not specified a variable in the "foreach" statement, $_ will take on each value in turn. For our example data, it would first be name=Nigel.
3. A search and replace; the first string (+) is replaced by a space, wherever it occurs.
4. This line looks for a % symbol followed by two characters: the brackets around the two periods mean that the characters will be saved as $1; the pack command is used to turn the hex code back into a single ASCII character. Once that's done, everything's decoded.
5. Now we chop up the information at the equals sign: to the left is the name of the field, and to the right is the info it contains.
6. Finally, we use Perl's associative arrays to store the data: by doing this, we don't have to alter the code for different forms.

The rest of your script can do whatever it likes after you have grabbed the input from the form: writing it to a disk file or creating a page, for instance. When you want to refer to data from the web page, you can do it simply by referring to, for example, $i{'name'}. And the rest of your script? That's up to you…

■ Starting this month, *Cutting Edge Net Answers* becomes *Hands On Internet*, with more comment and hands-on material. If there are any relevant subjects you'd like to see covered, please let us know.

PCW Contact

Nigel Whitfield is a freelance journalist, maintainer of several internet mailing lists and consultant to a number of non-profit organisations. He can be contacted via **internet@pcw.co.uk**.