Ergonomics of the down arrow

Ben Klemens

19 May 2009

[Part four of six]

[I wrote this in late 2005:] Google recently put out an RSS reader. It's pretty cute, and I personally have switched to it.

If you aren't familiar with RSS, then that is no matter here (it's a syndication system for web sites). The interesting feature of the reader for our purposes is that the J key will let you go down in the list of headlines. Yes, J, as in, uh, *jo down*. K, as in *kup* goes up in the list. There is absolutely nothing mnemonic about the J and K keys, but they *feel wonderful*. I assume you knows how to type properly, with hands on the home keys; I generally find my hands are on the home keys even when I'm just staring at the screen, and my hand doesn't need any help from my brain to find the little nubbin on the J key.

But that J key. It's the index finger of 90% of the world's dominant hand, and the keyboard is designed so that that index finger knows exactly where to rest. Moving down on the page is the most common operation, both in reading and even editing, so it makes complete ergonomic sense to attach this to the strongest finger of the strongest hand. Even the lefties will have no problem with it.

But it flies in the face of all mnemonics. Maybe you can come up with some word having to do with the process of scrolling down that begins with the letter J, but I've got nothin'. Nor could I think of a more efficient keymap.

I personally think the use of the J key is easy to learn because of its ergonomic delight. But it throws ease of initial use out the window—almost belligerently. You want to use the nifty hotkeys? Then RTFM.

An interface which works *against* intuition can be destructive, so if U went down and D went up, we'd have to write off the application as hopeless, but J doesn't work against anything. It's just a gesture.

Within a week of Google's RSS rollout, Bloglines, a competing RSS aggregation service, added a little header to its page: "You can now navigate through Bloglines with hotkeys[. . .]: j - next article k - previous article [. . .]"

Anybody familiar with the OpenOffice.org internals? bdamm (at) openoffice (dot) org will give you a hundred bucks to write code to have J move the cursor down a line (plus a handful of other keystrokes like K).

The war   Lest you think this J thing is some sort of recent meme, it all comes from vi, a text editor written in 1976. I am using a version of vi (named vim) to write this right now. Let's pause for a second and let that sink in: most programs have a shelf

life of about six months, and this guy wrote a program thirty years ago which is still in somewhat common use today. j goes down, k goes up, `{jfw` will go to the first instance of the letter w in your paragraph, and, since I can't stand seeing that unclosed open-bracket, I have to tell you that `}j%d%` will delete a parenthetical remark in the first line of the next paragraph. Which is all to show you that Mr. Joy, the author of vi, fell soundly on the efficiency side of the efficiency vs intuition scale—and that is why his text editor has survived for thirty years, and is being imitated by cutting-edge web services.

We sometimes like to write documents that actually have Js in them, and vi thus has modes: in editing mode, j goes down and d$ will delete the rest of the line; in insert mode, the j key puts a j on the screen, and typing d$ puts gibberish on the screen which quickly reminds you you're in the wrong mode.

There are two competitors to J. The first is the ctrl-D school, rooted in EMACS, written by a certain Mr. RM Stallman. EMACS's keymap is sort of like vi's, in that it's not particularly intuitive, but once you've learned it, you're done. However, it's a compromise along the efficiency vs intuition scale, because you don't need to deal with the unintuitive modes but reaching for the ctrl key all the time is not nearly as pleasant as twitching your index finger to hit the j key.[1] The EMACS vs vi war is a long-standing one, which is just silly, because they're of basically comparable efficiency. No, there are other schools that are a real drain on the economy, like the down-arrow school.

Let me take a paragraph or two to make this as clear as possible: the down-arrow school is a total failure when it comes to efficiency. On my screen right now, getting to the first w in the last paragraph via arrow keys is 27 keystrokes (using ctrl-arrow to go by word where possible). It's about three or four seconds for a single navigation. Do forty three-second navigations in a day and you're already up to nine hours in a work-year—a full work day a year just hitting the arrow key. You get to multiply by your wage to see what your company is spending per annum to facilitate ease of initial use. Even if it's one tap of the arrow key, your hands are already off the home keys; going off and on again is another half-second. If you do a hundred arrow-key navigations in a day (and if you're an office worker who does a lot of writing, you probably do closer to a thousand), that's another full work day a year just moving your right hand back and forth between the arrow keys and the home keys.

There is only one school that fails with such vehemence that it makes the down-arrow school look like Nirvana: the mouse school. In the mouse school, you take one hand—typically your dominant hand—off of the keyboard entirely, reaching to some part of the desk that is ergonomically suboptimal (because your keyboard is already in the optimal location). You position your hand on the mouse, and then move the cursor along the screen. It is an analog device, so aim and precision matter, meaning that some people simply do not have the eyesight and dexterity to use the mouse at all: try getting Aunt Myrtle to highlight the letter *i* in a font where that letter is one pixel wide. You guide the mouse to the pixels that are by the word you want to change, click, carefully drag, and return your hand to the keyboard. The entire process can easily take more than four or five seconds, just to position the cursor. And if you have to scroll through the document to find the point, that's easily ten seconds as prelude to a single edit.

---

[1]The joke is that EMACS stands for escape meta alt control shift.

The rabidness of the aforementioned text editor wars comes from the fact that text editing absorbs a huge amount of one's life. If you're like most office drones, most of your time at the computer is spent writing and editing plain text—and you're just one office drone; there are millions in the U.S.A. who are all operating computers basically identical to yours, using a down-arrow/mouse school text editor of some sort. Sure, there are people doing flashy data-slinging with big servers, but the bulk of computing is the literally billions of person hours per year spent editing text. Now multiply that half-second to move the right hand to the arrow key; at this scale, it adds up to millions of person-days per year spent on making that little twitch. With an entirely straight face, I can say that on the order of a billion dollars per year is spent on paying people to hit arrow keys.

When the programmer guys got together and wrote whatever it is you use to write your documents and navigate your web pages, they had all of the paradigms above at hand. Half of these guys are using EMACS or vi themselves. We get frustrated when we ask Mr. Computer Geek for help and he (always a boy, eh) comes back with over-everyone's-head exposition about just opening up regedt and doing a quick ctrl-f for HKEY {343-f2ea53e}. Less blatant but just as insidious is when Mr. Geek assumes you are an idiot. He knows that he knows more about PCs than you do, therefore you are dumb and wholly incapable of learning the reams of knowledge that he has compiled. I have been at many a workplace with IT departments that are stocked with such people; it's only some vestige of courtesy that keeps them from installing drool-guards on all the company keyboards.

Of course, the IT department is thinking about the worst-case users. But when was it ever efficient to force everybody in a several-hundred person organization to work with exactly the tools that the least-able could work with? You may have a legally blind worker at your workplace, but that doesn't mean that every computer in the building needs to operate exclusively at super-magnified resolution. A reasonable approach would be a system where you could select between the various schools of navigation. Most versions of vi let you do this (and EMACS allows ctrl-D, down-arrow, and a limited j-mode), but few down-arrow school programs include the wealth of editing keystrokes that those programs provide.

And so I take Google's j and k keys as a slight victory in a long battle against the forces of condescension. It's just two keys, a far cry from a word processor with a full vi keymap, but it's a sign that the guys who designed and programmed the system felt that it was more important to make usage efficient than to make it drool-proof. As such, it gives me hope that maybe the software of the future might focus on long-term efficiency over the quick sell.

Formatting and ergonomics Beyond editing, all this applies to formatting in Word too, because you have to use the mouse or an absurd amount of tabbing and arrowing to navigate the menus and dialog boxes to get to the option you want to change. For almost every step of the way, Word eagerly picks intuition over efficiency.

Of course, the most commonly-used features, like boldface, have their own ctrl-key combination, to at least save the user mouse and arrow-key inefficiency for the dozen most commonly-used operations. Also, you can use alt-F to access the File menu, alt-E

to use the Edit menu, et cetera.

But even having the few control-key combinations you do have creates problems, because there are only 26 control-letters to use. If they are taken up with the typesetting features of Word, then they can't be used for the plain old editing of text. EMACS and vi give the user fifty-odd keystrokes that edit text (I'm guessing because I couldn't possibly count them all); Word gives you cut, paste, copy, and that's about it. For every other editing task, you have to make do with the arrow keys. The majority of your time putting together a paper is spent writing and editing, so having so many keystrokes at your fingertips for formatting but almost none for editing is backward.

There is no place in Word's intuitive editing model for a key combination to delete a word at a time, to repeat the last edit, to jump to wherever you were last working, or to switch a lowercase letter to capital. But such keystrokes provide immense speed gains to users who have taken the time to learn them. But which do you do more often in a day: skip back to the beginning of a sentence, or switch to boldface?

One reason we have so many formatting commands is—once again—the lack of style sheets, which means that formatting is not produced by listing what you want the formatting to look like, but by applying it over and over again, which means that keystrokes to apply formatting are competing with editing keystrokes for frequency of use. It would be nice to have a dedicated editing program plus a separate dedicated formatting program, but Word's DOC format precludes this.

More on this next time.