Subject: Phil Agre's Mind
From: "pjk" <pjk@design.eng.yale.edu>
Date: 12 Jul 2000 18:47:39 -0400
From pjk@design.eng.yale.edu Wed Jul 12 18:50:28 2000

Dear Colleagues -

I think you know well that I value the comments of Phil Agre of UCLA in matters of Internet and culture. I’ve passed on a number of them through this list.

A mailing from Phil today gives a refreshing insight into how he evolved his analytical approach, starting from a college background in math and engineering. I think you’ll find it rewardingly provocative reading.

Favorite quote

> I finally comprehended the difference between critical thinking and its opposite. Technical people are not dumb, quite the contrary, but technical curricula rarely include critical thinking in the sense I have in mind. Critical thinking means that you can, so to speak, see your glasses. You can look at the world, or you can back up and look at the framework of concepts and assumptions and practices *through* which you look at the world.

All best, --Peter

-------------------------------------------------------
(from "RRE notes and recommendations, 7/12/00)

You may recall that I did my doctoral work in the field of artificial intelligence. I got into AI from adolescent motives; if you were a geek in the 1970s then that's what was cool. Having committed myself to a lengthy program of graduate study, however, I started to grow up. That's not to say that all AI people are immature, only that my own interest in the topic was the sort of thing one outgrows. But professional training has an incredible power to shape one's identity and thinking, and it is not easily outgrown.

One way to relate to AI, or to anything really, is the engineering way: if it can solve engineering problems then it's good. But that's not my way. I want to know if it's true. And not true in a logical or scientific sense but intuitively. I'm not very good at puzzles and problems. I don't focus well. I can chop logic with the best of them, but I don't like it. The thing is, logic is not that much of a
constraint. Ideas can be logical and nonsensical at the same time: it happens every day. So while I disapprove of illogic, I don't try to navigate by logic. I do big patterns. And the first big pattern that I wanted to discern was the underlying structure of AI. What *was* it about AI? Why did it seem so wrong? No incremental answer would do; the innumerable aspects of AI tended to reinforce one another, so that you couldn't change just one. You had to rethink the whole thing. You had to get the whole thing into your head. And this is what I do: I try to get the whole world into my head.

I've told aspects of this story before, for example in an intellectual memoir of my AI years entitled "Toward a critical technical practice" that you can find on my home page. Here I want to tell a smaller story, about how I became (relatively speaking, and in a small way) a better person through philosophy.

So here I was in the middle of the AI world -- not just hanging out there but totally dependent on the people if I expected to have a job once I graduated -- and yet, day by day, AI started to seem insane. This is also what I do: I get myself trapped inside of things that seem insane. Yet AI was all I knew. I was a math major in college, so I had never been compelled to learn much that was softer than electromagnetism. And I sure didn't have the morbid preoccupation with theorems that it would take to succeed in math. So I was on my own. Actually, not quite. Several other AI Lab students had the same bad attitude about AI and respect for real intellectual stuff that I did, so we were a club. That was my first mailing list, in fact. And we made contact with some critics of AI, who thought we were strange but helped us anyway. Still, there was no getting around it: when you're in that sort of hole, beyond a certain point you're on your own.

Although the AI Lab has always been capable of great strangeness, I would conjecture that my dissertation is one of the stranger documents ever accepted for a PhD at MIT. If you can print Postscript files you can still pull it down from the AI Lab publications Web site at:


After years of late nights scrawling endless stream-of-consciousness theorizing and example-working into my notebook, I wrote half of my dissertation in the last six weeks. There being no good reason not to graduate, I had rented a truck, moved all my belongings into it, moved out of my place in Boston, started paying rent on a place in San Francisco, parked the truck out in front of the AI Lab, camped out in my office, played records by Husker Du and Voivod at high volume, and wrote the stupid thing. I didn't even look for a job. When I got done, I did the paperwork, climbed in the truck, and drove to San Francisco. I made one stop along the way: at the gift shop of the Strategic Air Command museum in Nebraska, where they have these cool coffee mugs with the SAC logo on them. Freaked out, I moved my stuff into my new apartment, headed to the airport, flew to Paris, drove the most circuitous route I could find from Paris to Geneva to Lisbon, flew home, and beheld what I had done.

And there I was. I had no job, no money, and no clue what I wanted to do next. I was exhausted; I can't begin to tell you how exhausted I was. I would emerge from one layer of exhaustion, only to remember ten more layers of exhaustion that I hadn't even thought about. And I was angry. Not a good anger, mind you. Normally when people say, "I was angry", they mean to imply that they were justified. But it wasn't like that. It was more like, I'd been at war for years, and most of the war was in my own head. Only with the physical distance,
and having the dissertation safely in the library, could I even start
to get any perspective on it. To a rational person it would have
seemed like an impossible situation, and yet somehow I was certain
that I know what I am doing. I have no idea why. Yet somehow, as it
always does, the way forward materialized in front of me. I do not
know how I work this. And I do not defend it. It's just what happens.

I borrowed money from my former MIT officemate, who was raking it in
at Microsoft, and I got a consulting job through another friend in
the Valley. I visited with like-minded people at places like Xerox.
But for the most part I went to a bar -- the Vesuvio in North Beach.
I got a mineral water and lime, I went upstairs where no one would
bother me, and I resumed writing stream-of-consciousness theories
into my notebook, to see what would come out. I could never explain
all the intermediate steps that brought me to this, but I decided
to try an experiment. In graduate school I had been taken with the
method of Derrida. Never mind all of the American literary critics
who have gone to extremes with deconstruction. I was still basically
an engineer, and to my engineer's mind Derrida's method of applying
the techniques of literary close-reading to philosophical texts made
all kinds of sense. You don't understand what Derrida's getting at
if all you've read is the comic book or op-ed version of his strange
conclusions. It's a discipline he's following, a serious one, and
you only understand it if you follow it step by step. My experiment
was to apply these same techniques to technical texts.

I got out a photocopy of the first chapter of one of the founding
texts of AI, Miller, Galanter, and Pribram's "Plans and the Structure
of Behavior" (1960), and I read the first paragraph. Just the first
paragraph. I read it word-by-word, writing endless speculative notes
about the metaphors and rhetoric and prosody and so forth of every
word. I actually wrote that first paragraph out on a 3-by-5 card
and carried it with me wherever I went. Analyzing that paragraph
took weeks. But a picture emerged. The paragraph, it turned out,
was really about hypnosis. The paragraph was trying to hypnotize
its reader, and it was trying to merge the authors and reader rather
than establish a dialogue among them. This was not just one of those
arbitrary just-so stories that bad literary critics tell when they
get going into a high-theory interpretive frenzy. I hate that stuff.
No: this was a tight, logical argument about the fancy stuff -- the
"economy", as Derrida would say -- that was going on beneath the
surface of the text. You can read this argument on my Web site:

http://dlis.gseis.ucla.edu/people/pagre/portents.html

I was deeply impressed by this. I figured that I could make a career
out of such readings, and over the next year I wrote a long paper that
interprets the Miller, Galanter, and Pribram text and a few others.
I was concerned with formalization, which is the process by which
mathematizing science and technology fields translate natural language
into mathematics. What happens to language during that translation
process? An awful lot, in fact, and with close reading we can watch
language go through characteristic deformations that tell us an awful
lot about why so many technical fields are both dangerous nonsense and
powerful ways of looking at the world.

I never published this long paper, but I drew confidence from it.
Parts of it show up in my subsequent papers about AI, all available
on my Web site, and in my book. But its most important effect was
inside of me. I finally comprehended the difference between critical
thinking and its opposite. Technical people are not dumb, quite the
contrary, but technical curricula rarely include critical thinking in
the sense I have in mind. Critical thinking means that you can, so to
speak, see your glasses. You can look at the world, or you can back up and look at the framework of concepts and assumptions and practices *through* which you look at the world. Every such framework edits the world in some way; every such framework has its biases. And no matter how carefully you think you define your words, most of your framework of concepts and assumptions and practices for looking at the world will be inherited from a long disciplinary and cultural tradition. If you can't see your glasses then you will have tunnel vision your whole life. Yet you probably won't even notice, because your ways of looking at the world also define what counts as success, as progress, as a research result, and so on. Not that critical thinking makes you omniscient: you're still wearing glasses even when you're looking at your glasses. This (and not any sort of silly idealism) is what Derrida means when he says that a text has no outside. But through scholarship and analysis you can do a lot better than just stumbling along with the glasses you got in school.

So I found a way of relating to the technical texts that had struck me as wrong-headed: they were data, objects of investigation. I could read them for all of the baroque stuff that was going on within them. And let me tell you, there is a lot of baroque stuff going on inside your average AI text. That's not because the author of said text is dishonest or crazy; that person is not trying to be perverse, but has been socialized into a way of talking and thinking, and has not been socialized into any critical practice for understanding what's really going on. The average AI author is not try to say crazy stuff; it is more accurate to say that he or she is trying *not* to say crazy stuff, but that the inherited discursive forms of AI won't cooperate. I see the crazy stuff, or some of it, and I thrash myself trying to figure out how to explain it. It's like trying to explain the ocean to a fish; the concept simply makes no sense to someone who lives inside the thing I am trying to explain.

That's obviously no good. It doesn't help anyone. Fortunately, however, and through no conscious planning on my part, the situation evolved further. I somehow no longer experience a strong, dichotomous opposition between two relations to a text: taking it seriously or holding it at arm's length as an object of critical analysis. That strong distinction, it turns out, is an artefact of my having started with the worst possible case, the most ambitious of all technical disciplines, AI. The great virtue of the interpretive social sciences is that whole generations of smart people have applied a critical sensibility to their own work, "reflexively" as they say, so that whole generations of students have grown up being socialized into a relatively conscious relation to their own language and assumptions and practices. The great Foucauldian revolution of recent years has destroyed much of this progress, I have to admit, at least in a broad swathe of social theory, since the followers of Foucault think that discourses are identical with the real world, as opposed to being optional and contingent ways of looking at it. Nonetheless, the older practices are still there, still very much alive in other precincts, and despite their reputation for negativity their habit of self-suspicion can bring a tolerance for the tunnel vision of others. So it is with me: AI may be sunk to its eyeballs in unexamined assumptions, but so am I.