

Sandy: I guess I'm a strange sort of advocate for machine intelligence. To some degree I straddle the fence. I think that machines won't really be intelligent in a humanlike way until they have something like your biological wetness or slipperiness to them. I don't mean literally wet-the slipperiness could be in the software. But biological seeming or not, intelligent machines will in any case be machines. We will have designed them, built them-or grown them! We'll understand how they work-at least in some sense. Possibly no one person will really understand them, but collectively we will know how they work.

Pat: It sounds like you want to have your cake and eat it too. I mean, you want to have people able to build intelligent machines and yet at the same time have some of the mystery of mind remain.

Sandy: You're absolutely right-and I think that's what will happen. When *real* artificial intelligence comes-

Pat: Now there's a nice contradiction in terms!

Sandy: Touche! Well, anyway, when it comes, it will be mechanical and yet at the same time organic. It will have that same astonishing flexibility that we see in life's mechanisms. And when I say mechanisms, I mean mechanisms. DNA and enzymes and so on really are mechanical and rigid and reliable. Wouldn't you agree, Pat?

Pat: Sure! But when they work together, a lot of unexpected things happen. There are so many complexities and rich modes of behavior that all that mechanicalness adds up to something very fluid.

Sandy: For me, it's an almost unimaginable transition from the mechanical level of molecules to the living level of cells. But it's that exposure to biology that convinces me that people are machines. That thought makes me

uncomfortable in some ways, but in other ways it is exhilarating.

Chris: I have one nagging question If people are machines, how come it's so hard to convince them of the fact? Surely a machine ought to be able to recognize its own machinehood!

Sandy: It's an interesting question. You have to allow for emotional factors here. To be told you're a machine is, in a way, to be told that you're nothing more than your physical parts, and it brings you face to face with your own vulnerability, destructibility, and, ultimately, your mortality. That's something nobody finds easy to face. But beyond this emotional objection, to see yourself as a machine, you have to "unadopt" the intentional stance you've grown up taking toward yourself-you have to jump all the way from the level where the complex lifelike activities take place to the bottommost mechanical level where ribosomes chug along RNA strands, for instance. But there are so many intermediate layers that they act as a shield, and the mechanical quality way down there becomes almost invisible. I think that when intelligent machines come around, that's how they will seem to us-and to themselves! Their mechanicalness will be buried so deep that they'll seem to be alive and conscious-just as we seem alive and conscious

Chris: You're baiting me! But I'm not going to bite.

Pat: I once heard a funny idea about what will happen when we eventually have intelligent machines. When we try to implant that intelligence into devices we'd like to control, their behavior won't be so predictable.

Sandy: They'll have a quirky little "flame" inside, maybe?

Pat: Maybe.

Chris: And what's so funny about that?

Pat: Well, think of military missiles. The more sophisticated their target-tracking computers get, according to this idea, the less predictably they will function. Eventually, you'll have missiles that will decide they are pacifists and will turn around and go home and land quietly without blowing up. We could even have "smart bullets" that turn around in midflight because they don't want to commit suicide!

Sandy: What a nice vision!

Chris: I'm very skeptical about all this. Still, Sandy, I'd like to hear your predictions about when intelligent machines will come to be.

Sandy: It won't be for a long time, probably, that we'll see anything remotely resembling the level of human intelligence. It rests on too awesomely complicated a substrate-the brain-for us to be able to duplicate it in the foreseeable future. Anyhow, that's my opinion.

Pat: Do you think a program will ever pass the Turing Test?

Sandy: That's a pretty hard question. I guess there are various degrees of passing such a test, when you come down to it. It's not black and white. First of all, it depends on who the interrogator is. A simpleton might be totally taken in by some programs today. But secondly, it depends on how deeply you are allowed to probe.

Pat: You could have a range of Turing Tests-one-minute versions, five-minute versions, hour-long versions, and so forth. Wouldn't it be interesting if some official organization sponsored a periodic competition, like the annual computer-chess championships, for programs to try to pass the Turing Test?

Chris: The program that lasted the longest against some panel of distinguished judges

would be the winner. Perhaps there could be a big prize for the first program that fools a famous judge for, say, ten minutes.

Pat: A prize for the program, or for its author.

Chris: For the program, of course!

Pat: That's ridiculous! What would a program do with a prize?

Chris: Come now, Pat. If a program's human enough to fool the judges, don't you think it's human enough to enjoy the prize? That's precisely the threshold where it, rather than its creators, deserves the credit, and the rewards. Wouldn't you agree?

Pat: Yeah, yeah-especially if the prize is an evening out on the town, dancing with the interrogators!

Sandy: I'd certainly like to see something like that established. I think it could be hilarious to watch the first programs flop pathetically!

Pat: You're pretty skeptical for an AI advocate, aren't you? Well, do you think any computer program today could pass a five minute Turing Test, given a sophisticated interrogator?

Sandy: I seriously doubt it. It's partly because no one is really working at it explicitly. I should mention, though, that there is one program whose inventors claim it has already passed a rudimentary version of the Turing Test. It is called "Parry," and in a series of remotely conducted interviews, it fooled several psychiatrists who were told they were talking to either a computer or a paranoid patient. This was an improvement over an earlier version, in which psychiatrists were simply handed transcripts of short interviews and asked to determine which ones were with a genuine paranoid and which ones were with a computer simulation.

Pat: You mean they didn't have the chance to ask any questions? That's a severe handicap-and it doesn't seem in the spirit of the Turing Test. Imagine someone trying to tell which sex I belong to, just by reading a transcript of a few remarks by me. It might be very hard! I'm glad the procedure has been improved.

Chris: How do you get a computer to act like a paranoid?

Sandy: Now just a moment-I didn't say it does act like a paranoid, only that some psychiatrists, under unusual circumstances, thought so. One of the things that bothered me about this pseudo Turing Test is the way Parry works. "He," as the people who designed it call it, acts like a paranoid in that "he" gets abruptly defensive and veers away from undesirable topics in the conversation. In effect, Parry maintains strict control so that no one can truly probe "him." For reasons like this, simulating a paranoid is a whole lot easier than simulating a normal person.

Pat: I wouldn't doubt that. It reminds me of the joke about the easiest kind of human being for a computer program to simulate.

Chris: What is that?

Pat: A catatonic patient-they just sit and do nothing at all for days on end. Even I could write a computer program to do that!

Sandy: An interesting thing about Parry is that it creates no sentences on its own-it merely selects from a huge repertoire of canned sentences the one that in some sense responds best to the input sentence.

Pat: Amazing. But that would probably be impossible on a larger scale, wouldn't it?

Sandy: You better believe it (to use a canned remark)! Actually, this is something that's really not appreciated enough. The number of

sentences you'd need to store in order to be able to respond in a normal way to all possible turns that a conversation could take is more than astronomical-it's really unimaginable. And they would have to be so intricately indexed, for retrieval... Anybody who thinks that somehow a program could be rigged up just to pull sentences out of storage like records in a jukebox, and that this program could pass the Turing Test, hasn't thought very hard about it. The funny part is that it is just this kind of unrealizable "parrot program" that most critics of artificial intelligence cite, when they argue against the concept of the Turing Test. Instead of imagining a truly intelligent machine, they want you to envision a gigantic, lumbering robot that intones canned sentences in a dull monotone. They set up the imagery in a contradictory way. They manage to convince you that you could see through to its mechanical level with ease, even as it is simultaneously performing tasks that we think of as fluid, intelligent processes. Then the critics say, "You see! A machine could pass the Turing Test and yet it would still be just a mechanical device, not intelligent at all." I see things almost the opposite way. If I were shown a machine that can do things that I can do-I mean pass the Turing Test-then, instead of feeling insulted or threatened, I'd chime in with philosopher Raymond Smullyan and say, "How wonderful machines are!"

Chris: If you could ask a computer just one question in the Turing Test, what would it be?

Sandy: Ummm . . .

Pat: How about this: "If you could ask a computer just one question in the Turing Test, what would it be?"

END.

Post Scriptum: "Nicolai"

In 1983, I had the most delightful experience of getting to know a small group of extremely enthusiastic and original students at the University of Kansas in Lawrence. These students, about thirty in number, had been drawn together by Zamir Bavel, a professor in the Computer Science Department, who had organized a seminar on my book *Gödel, Escher, Bach*. He contacted me and asked me if there was any chance I could come to Lawrence and get together with his students. Something about his way of describing what was going on convinced me that this was a very unusual group and that it would be worth my while to try it out. I therefore made a visit to Kansas and got to know both Zamir and his group. All my expectations were met and surpassed. The students were full of ideas and warmth and made me feel very much at home.

The first trip was so successful that I decided to do it again a couple of months later. This time they threw an informal party at an apartment a few of them shared. Zamir had forewarned me that they were hoping to give me a demonstration of something that had already been done in a recent class meeting. It seems that the question of whether computers could ever think had arisen, and most of the group members had taken a negative stand on the issue. Rod Ogborn, the student who had been leading the discussion, had asked the class if they would consider any of the following programs intelligent:

1. A program that could pass a course in beginning programming (i.e., that could take informal descriptions of tasks and turn them into good working programs);
2. A program that could act like a psychotherapist (Rod gave sample dialogues with the famous "Doctor" program, also known as "ELIZA," by Joseph Weizenbaum);

3. A program called "Boris," written at Yale by Michael Dyer, that could read stories in a limited domain and answer questions about the situation which required filling in many unstated assumptions, and making inferences of many sorts based on them.

The class had come down on the "no" side of all three of these cases, although they got progressively harder. So Rod, to show the class how difficult this decision might be if they were really *faced* with a conversational program, managed to get a hookup over the phone lines with a natural-language program called "Nicolai" that had been developed over the last few years by the Army at nearby Fort Leavenworth. Thanks to some connections that Rod had, the class was able to gain access to an unclassified version of Nicolai and to interact with it for two or three hours. At the end of those hours, they then reconsidered the question of whether a computer might be able to think. Still, only one student was willing to consider Nicolai intelligent, and even that student reserved the right to switch sides if more information came in. About half the others were noncommittal, and the rest were unwilling, under any circumstances, to call Nicolai intelligent. There was no doubt that Rod's demonstration had been effective, though, and the class discussion had been one of the most lively.

Zamir told me all of this on our drive into Lawrence from the Kansas City airport, and he explained that the group had been so stimulated by this experience that they were hoping to get reconnected to Nicolai over the phone lines, and to let me try it out during the party. I thought it sounded quite amusing, and since I have tried out and watched a lot of natural-language programs in my time, I thought I would have an easy time coming up with good probes into the weaknesses of Nicolai. "Besides," I thought (somewhat chauvinistically, "how good can a program be that is being developed by the Army in an out-

of-the-way place like Fort Leavenworth? To reveal its mechanicalness ought to be a piece of cake for me!" So I was looking forward to the challenge.

When we first arrived at the party, everyone was drinking beer and munching on various chips and dips and talking excitedly about what had happened in the previous class meeting with Nicolai. There was some consternation, though, because it seemed that the connection to Nicolai was hard to reestablish. It had to go through a computer at the University of Kansas, and thence to Fort Leavenworth, all of which meant that it was a slow and tenuous link. But after about 45 minutes of trying to set up the link, someone announced with pleasure that Nicolai was responding. One of the students was sitting at the terminal and typing in various simple questions and getting back simple answers. I watched over his shoulder, and felt the itch to sit down and try to unmask Nicolai for what "he" really was. After a couple of minutes, the student relinquished the terminal and I got my chance, with a small group of people looking over my shoulder as I typed. Nicolai's answers were quite slow in coming over the phone lines, and the wait was made all the more suspenseful by the fact that they were transmitted only in full lines of text, so there were some pretty long delays between successive lines of response. But that didn't seriously dampen my enthusiasm for trying to quickly poke holes in the program's facade of understanding.

Herewith follows a somewhat abridged but accurate transcript of the "conversation" I had with Nicolai over the next hour or so. For convenience's sake, what Nicolai typed is shown in capitals, and my inner thoughts at the time (as best I can reconstruct them now) are in parentheses.

> What would you like to talk about?

(Let's start out easy.)

I CAN TALK ABOUT ANYTHING.

(Yeah, yeah--tell it to the Marines!)

> Is there anything you specialize in?

WHAT DOES "SPECIALIZE" MEAN? PLEASE USE A COMPLETE SENTENCE.

(For a supposedly smart program, Nicolai doesn't seem to have a very big vocabulary.)

> "Specialize" means "to be good at."

(I wonder if Nicolai likes those quotation marks, or if they will throw it for a loop.)

I UNDERSTAND.

(Izzatso?)

> When were you written?

LAST COMPILATION: "4/27/83 19.247"

(Oh, look at that stupid, rigid answer, complete with pointless marks and that extra, unexplained number at the end. I'm exposing Nicolai pretty well, so far....)

> What language is Nicolai written in?

(I think I'll see if it knows its own name and recognizes third-person references to itself.)

NICOLAI IS WRITTEN MOSTLY IN ADA.

(Hmmm . . . not bad. It handled both "you" and "Nicolai" when I talked about it. Those Army programmers aren't so bad!)

> What are feet?

(Choose an ambiguous word, see what it comes up with....)

FEET ARE 12.0 INCHES.

(Wow, weird! A person would never put it that way. They'd say, "A foot is a unit of length, twelve inches long," or something like that. And anyway, feet are primarily parts of the body, secondarily units of length. Why didn't it even seem to recognize that ambiguity? Oh well, I guess that kind of dumbness is the kind of thing I wanted to expose. Score one for me.)

> What are ears?

(Why not try another ambiguous word? After all, I did well on the last one.)

EARS ARE AUDITORY ORGANS FOUND ON ANIMALS.

(This is really funny! Why were feet not thought of primarily as body parts, then? This body-parts stuff seems to give me lots of points-I think I'll try it again.)

> What are arms?

THAT INFORMATION IS CLASSIFIED.

(Oh no! That is too stupid for words! Or is this somebody's idea of a joke?)

> I like to understand up without seventeen clams.

(Let's toss some nonsense at Nicolai and see what it does.)

WHAT DO CLAMS HAVE TO DO WITH IT?

(Oh come on, Nicolai! Do you mean you understood all the rest, except for the clams bit? You are really dumb!)

> Why are you so dumb?

I AM A SMART PROGRAM.

(Well, that's kind of touching, actually. A little self-defense is appealing, in a way.)

> You are pretty confident of yourself, aren't you?

WHAT DOES "PRETTY CONFIDENT" MEAN?

(Oh yeah, I should have known--pretty- as an adverb might be too colloquial.)

> Did you understand up without my last sentence?

(I figure, since earlier it didn't complain about my phrase "understand up," let's see what it does with it this time.)

YOUR SENTENCE IS NONSENSE.

(Oh really? Well, then, why wasn't the other one with that same phrase in it also considered nonsense?)

> Do you enjoy nonsense?

IF YOU INSIST ON TALKING NONSENSE, I HAVE OTHER TASKS TO PERFORM.

(A strange response. Who ever insisted on talking nonsense? I was just asking a simple question. Oh well, let me try out another angle.)

> What is your opinion of M. C. Escher?

M. C. ESCHER, ARTIST, TWENTIETH CENTURY

(Oh, not again-another vacuous dictionary look-up. Can't it do better than that, especially when I've asked it a probing question about taste? Let's see how it handles a complaint....)

> You did not answer my question.

WHICH QUESTION DID I NOT ANSWER?

(Hey, this Nicolai is a pretty slick character. Not so different from Parry. I must say, I am really pretty impressed by the robustness of this program. I mean, we have covered a lot of ground so far, and at least it makes "pseudo-sensible answers" most of the time. Not at all bad!)

> Do you own any money?

I DO NOT OWN ANYTHING.

(Fair enough.)

> Have you seen the latest issue of "Scientific American"?

(Uh-oh-once again, I don't know if I'm going to help it or confuse it by putting in those quotes. Let's see.)

I DO NOT HAVE ANY "SCIENTIFIC AMERICAN."

(That sounds awkward. Very computerish. I doubt if it knows what I meant. Oh well.)

YOU'VE BEEN TALKING TO ME FOR QUITE A WHILE. WHAT IS YOUR NAME?

(Hmm.... That was an unexpected remark by Nicolai. Should I answer it? Sure, why not?)

> My name is Doug.

DO YOU HAVE A LAST NAME?

(Inquisitive little devil, isn't it?)

> Of course I do.

(Ask a stupid question, get a stupid answer. Let's see how it fields this one.)

WOULD YOU CARE TO TELL ME WHAT YOUR LAST NAME IS?

(Damn, it just totally ignored the tone of my remark. And why is it being so polite, suddenly?)

> Only if you'll tell me yours.

(That ought to challenge it.)

I DO NOT HAVE A LAST NAME.

(I don't seem to be able to foil its defenses very easily. It always walks around me. Those programmers sure anticipated a lot of ideas. I guess they must have debugged Nicolai for years, watching typical dialogues with people who just come in off the street. For a dumb program, Nicolai is pretty smart!)

> Well then, I won't tell you mine.

THAT'S NOT BEING VERY FRIENDLY.

(You know, sometimes, I hate to admit it this stupid program really does seem to understand something!)

> I'm sorry. How would you like me to give you a last name?

IF YOU WOULD LIKE TO DO THAT, THAT WOULD BE FINE WITH ME.

(Now that is obviously a canned sentence, or at least the latter half must certainly be. Score one more point for me. And yet I'm not doing nearly as well as I'd hoped I would be by now. What's wrong?)

> Your last name is "Rimsky-Korsakov."

THANK YOU.

(Well, that's a strange kind of thing to thank somebody for. Just goes to show how the rigidity of programs makes them say the strangest things.)

> Don't mention it.

(I'll play along with this game for a second.)

WHAT IS YOUR LAST NAME?

(My god, it remembered, after all this while, that as soon as it had told me its last name, I was going to tell it mine. Well, I'd better make good on that promise.)

> My last name is "Hofstadter".

ARE YOU "(32) DOUGLAS R. HOFSTADTER"?

(Now why, in the midst of so much elegance, does it suddenly act so crude, putting my name in quotes and including that silly extraneous index number? I can't for the life of me figure out why this program is so mixed in quality.)

> One and the same.

(Try out a strange idiomatic phrase.)

(83) "PARSING ERROR IN PREVIOUS INPUT"

(Aha-fooled it again. All of a sudden, a whole lot of mechanicalness is showing through.)

> How many letters are there in your last name?

(I think I'll see whether it can figure out ambiguous words by context.)

14 PLUS A HYPHEN.

(Good. Now I'll let "letters" switch meaning. What will it do?)

> How many letters are there in an ordinary envelope?

18.

(Oh no! It didn't recognize the shift! And it merely counted the number of letters in "an ordinary envelope"! But if that's the case, then how come in the earlier case it didn't count the letters in the phrase "your last name," but instead retrieved the referent? There's something so inconsistent and uneven about this program. It just goes to show how far we have to go before mechanical intelligence bears any resemblance to human fluidity.)

ARE YOU WORKING ON A NEW BOOK?

(I wonder what will happen if I simply sidestep that question. Anyway, I'd like to probe its ability to deal with humor.)

> Tell me a joke.

I WANT A COOOKIE.

(What kind of a joke is that? Hey, it spelled "cookie" wrong! Hmm . . . is that its joke? Or what is going on here?)

> By the way, I am sort of working on a new book. Care to hear about it?

(Can it handle weird phrases like "sort of"? And questions with no subject?)

NO, BUT IF YOU WOULD GIVE ME A COOKIE, I WOULD GLADLY PAY YOU TUESDAY.

(That sounds like a familiar line. Have I heard it somewhere before? Is that supposed to be the punch line of its "joke"?)

> Tell me another one.

HOW MANY COMPUTER PROGAMMERS DOES IT TAKE TO SCREW IN A LIGHT BULB?

(Okay, I see. It's got a library of jokes and riddles, and it just fishes them out at random. I'll play along.)

> 4.

INCORRECT. THAT IS A HARDWARE PROBLEM.

(Okay, so it told its joke. But I bet it has no idea what any of the terms in it mean, such as "programmer." - Hey, it left an 'r' out of "programmer"! I guess whoever input the joke misspelled the word. So let's see what happens when I ask about it...)

> What is a computer programmer?

WOULD YOU LIKE ME TO EXPLAIN THE JOKE?

(Hey, damn it, it turned the tables on me! Does it think I don't know what a "programmer" (or a programmer) is? I don't like this slipperiness. Every time I try to get it to act really dumb, it somehow slips out of it. Strange.... All right, let me probe some more about its understanding of the terms in that joke. How about a potentially off-color word?)

> What does "to screw" mean?

IN WHAT CONTEXT?

(Aha, at least it knows it's an ambiguous term.)

> In bed.

TO SCREW IN BED MEANS TO "HAVE SEX."

(Okay, fine, it has some definitions stored according to context. I guess this word must come up all the time in conversations. But can it go any further than that definition?)

> What is having sex all about?

I CAN RECOMMEND AN EXCELLENT BOOK ON THAT SUBJECT.

(What a great stock answer! One point for the Army programmers!)

> Tell me, tell me!

THE BOOK IS CALLED "GODEL, ESCHER, BACH."

(Huh? Say, wait a minute, this is too weird. I'm not so sure this is really a program. There's too damn much cleverness in these weird responses. I can't imagine bizarre mechanical devices coming up with this much coherence, even if it is completely off the wall, at times.)

> I like to call it "GEB" myself.

THAT IS VERY NICE, DOUG. WHAT ELSE DO YOU CALL "GEB"?

(All right, the jig's up. What's going on here?)

At this point, everybody started laughing. They had all been trying hard to keep from doing so the whole time, but at this point, they couldn't contain themselves. It was time to reveal to me what had been going on. They took me downstairs and showed me that three of the students - Roy Leban, Howard Darsche, and Garrick Stangle - had been collectively playing Nicolai. There was in reality no such program, and there hadn't ever been one. (In retrospect, I am reminded of the famous French mathematician Nicolas Bourbaki - a hypothetical person, actually an amalgam of over a dozen eminent mathematicians writing under that collective pen name.) There had indeed been a similar demonstration for the class a few days earlier, and the class, like me, had been taken in for a long time. In my case, Roy, Howard, and Garrick had worked very hard to give the impression of mechanicalness by spewing back "parsing error" and other indications of rigidity, and also by sending what looked very much like canned phrases from time to time. That way they could keep sophisticates like me believing that there was a program behind it all. Only by that point I was beginning to wonder just how sophisticated I really was.

The marvelous thing about this game is that it was, in many ways, a Turing Test in reverse: a group of human beings masquerading as a program, trying to act mechanical enough that I would believe it really was one. Hugh Kenner has written a book called *The Counterfeiters* about the perennial human fascination with such compounded role flips. A typical example is Delibes's ballet *Coppelia*, in which human dancers imitate life-sized dolls stiffly imitating people. What is amusing is how Nicolai's occasional crudeness was just enough to keep me convinced it was mechanical. Its "willingness" to talk about itself, combined with its obvious limitations along those lines (its clumsy revelation of when it was last compiled, for instance), helped establish the illusion very strongly.

In retrospect, I am quite amazed at how much genuine intelligence I was willing to accept as somehow having been implanted in the program. I had been sucked into the notion that there really must be a serious natural-language effort going on at Fort Leavenworth, and that there had been a very large data base developed, including all sorts of random information: a dictionary, a catalog containing names of miscellaneous people, some jokes, lots of canned phrases to use in difficult situations, some self-knowledge, a crude ability to use key words in a phrase when it can't parse it exactly, some heuristics for deciding when nonsense is being foisted on it, some deductive capabilities, and on and on. In hindsight, it is clear that I was willing to accept a huge amount of fluidity as achievable in this day and age simply by putting together a large bag of isolated tricks-kludges and hacks, as they say.

Roy Leban, one of the three inside Nicolai's mind, wrote the following about the experience of being at the other end of the exchange:

Nicolai was a split personality. The three of us (as well as many kibitzers) argued about

practically every response. Each of us had a strong preconceived notion about what (or who) Nicolai should be. For example, I felt that certain things (such as "Douglas R. Hofstadter") should be in quotation marks, and that feet should not be 12 inches, but 12.0. Howard had a tendency for rather flip answers. It was he who suggested the "classified" response to the "arms" question. And somehow, when he suggested it, we all knew it was right.

Several times during our conversation, I felt quite amazed at how fluently Nicolai was able to deal with things I was bringing up, but each time I could postulate some not *too* sophisticated mechanical underpinning that would allow that particular thing to happen. As a strong skeptic of true fluidity in machines at this time, I kept on trying to come up with rationalizations for the fact that this program was doing so well. My conclusion was that it was a very vast and quite sophisticated bag of tricks, no one of which was terribly complex. But after a while, it just became too much to believe. Furthermore, the mixture of crudity and subtlety became harder and harder to swallow, as well.

My strategy had been, in essence, to use spot checks all over the map: to try to probe it in all sorts of ways rather than to get sucked into some topic of its own choice, where it could steer the conversation. Daniel Dennett, in a paper on the depth of the Turing Test, likens this technique to a strategy taught to American soldiers in World War II for telling German spies from genuine Yankees.

The idea was that even if a young man spoke absolutely fluent American-sounding English, you could trip him up by asking him things that any boy growing up in those days would be expected to know, such as "What is the name of Mickey Mouse's girlfriend?" or "Who won the World Series in 1937?" This expands the domain of knowledge necessary from just the

language itself to the entire culture-and the amazing thing is that just a few well-placed questions can unmask a fraud in a very brief time-or so it would seem.

The problem is, what do you do if the person is extremely sharp, and when asked about Minnie Mouse, responds in some creative way, such as, "Hah! She ain't no *girlfriend*-she's a *mouse*!"? The point is that even with these trick probes that should ferret out frauds very swiftly, there can be clever defensive counter-maneuvers, and you can't be sure of getting to the bottom of things in a very brief time.

It seems that a few days earlier, the class had collectively gone through something similar to what I had just gone through, with one major difference. Howard Darsche, who had impersonated (if I may use that peculiar choice of words!) Nicolai, in the first run through, simply had acted himself, without trying to feign mechanicalness in any way. When asked what color the sky was, he replied, "In daylight or at night?" and when told "At night," he replied, "Dark purple with stars." He got increasingly poetic and creative in his responses to the class, but no one grew suspicious that this Nicolai was a fraud. At some point, Rod Osborn simply had to stop the demonstration and type on the screen, "Okay, Howard, you can come in now." Zamir (who was not in cahoots with Rod and his team) was the only one who had some reluctance in accepting this performance as that of a genuine program, and he had kept silent until the end, when he voiced a muted skepticism.

Zamir summarizes this dramatic demonstration by saying that his class was willing to view *anything on a video terminal* as mechanically produced, no matter how sophisticated, insightful, or poetic an utterance it might be. They might find it interesting and even surprising, but they would find some way to discount those qualities. Why was this the case? How could they do this for so long? And

why did I fall for the same kind of thing?

In interacting with me, Nicolai had seemed to waver between crude mechanicalness and subtle flexibility, an oscillation I had found most puzzling and somewhat disturbing. But I was still taken in for a very long time. It seems that, even armed with spot checks and quite a bit of linguistic sophistication and skepticism, unsuspecting humans can have the wool pulled over their eyes for a good while. This was the humble pie I ate in this remarkable reverse Turing Test, and I will always savor its taste and remember Nicolai with great fondness.

Alan Turing, in his article, indicated that his "Imitation Game" test should take place through some sort of remote Teletype linkup, but one thing he did not indicate explicitly was at what grain size the messages would be transmitted. By that, I mean that he did not say whether the messages should be transmitted as intact wholes, or line-by-line, word by word, or keystroke by keystroke. Although I don't think it matters for the Turing Test in any *fundamental* sense, I do think that which type of "window" you view another language-using being through has a definite bearing on how *quickly* you can make inferences about that being. Clearly, the most revealing of these possibilities is that of watching the other "person" operate at the keystroke level.

On most multi-user computer systems, there are various ways for different users to communicate with each other, and these ways reflect different levels of urgency. The slowest one is generally the "mail" facility, through which you can send another user an arbitrarily long piece of text, just like a letter in an envelope. When it arrives, it will be placed in the users "mailbox," to be read at their leisure. A faster style of communicating is called, on UNIX systems, "write." When this is invoked, a direct communications link is set up between you and the person you are trying to reach (provided they are logged on). If they accept

your link, then any full line typed by either of you will be instantly transmitted or printed on the other party's screen-where a lineup is signaled by your hitting the carriage-return key. This is essentially what the Nicolai team used in communicating with me over the Kansas computer. Their irregular typing rhythm and any errors they might have made were completely concealed from me this way, since all I saw was a sequence of completely polished lines (with the two spelling errors-"coookie" and "progammer," which I was willing to excuse because Nicolai generated them in a "joke" context).

The most revealing mode is what, on UNIX, is called "talk." In this mode, every single keystroke is revealed. You make an error, you are exposed. For some people, this is too much like living in a glass house, and they prefer the shielding afforded by "write." For my part, I like living dangerously. Let the mistakes fly! In computer-mediated conversations with my friends, I always opt for "talk." I have been amused to watch their "talk" styles and my own slowly evolve to relatively stable states.

When we in the Indiana University Computer Science Department first began using the "talk" facility, we were all somewhat paranoid about making errors, and we would compulsively fix any error that we made. By this I mean that we would backspace and retype the character. The effect on the screen of hitting the backspace key repeatedly is that you see the most recently typed characters getting eaten up, one by one, right to left, and if necessary, the previous line and ones above it will get eaten backward as well. Once you have erased the offending mistakes, you simply resume typing forward. This is how errors are corrected. We all began in this finicky way, feeling ashamed to let anything flawed remain "in print" so to speak, visible to others' eyes. But gradually we overcame that sense of shame, realizing that a typo sitting on a screen is not quite so deathless as one sitting on a page in a book.

Still, I found that some people just let things go more easily than others. For instance, by the length of the delay after a typo is made, you can tell just how much its creator is hesitating in wondering whether to correct it. Hesitations of a fraction of a second are very noticeable, and are part of a person's style. Even if a typo is left uncorrected, you can easily spot someone's vacillations about whether or not to fix it.

The counterparts of these things exist on many levels of such exchanges. There are the levels of *word choice* (for instance, some people who don't mind having their typos on display will often backtrack and get rid of *words* they now repudiate, *sentence-structure* choice, *idea* choice, and higher. Hesitations and repairs or restarts are very common. I find nothing so annoying as someone who has gotten an idea expressed just fine in one way, and who then erases it all on the screen before your eyes and proceeds to compose it anew, as if one way of suggesting getting together for dinner at Pagliai's at six were markedly superior to another!

There are ways of exploiting erasure in "talk" mode for the purposes of humor. Don Byrd and I, when "talking," would often make elaborate jokes exploiting the medium in various ways. One of his, I recall vividly, was when he hurled a nasty insult onto the screen and then swiftly erased it, replacing it by a sweetly worded compliment, which remained for posterity to see-at least for another minute or so. One of our great discoveries was that some "arrow" keys allowed us to move all over the screen, and thus to go many lines up in the conversation and edit earlier remarks by either of us. This allowed some fine jokes to be made.

One hallmark of one's "talk" style is one's willingness to use abbreviations. This is correlated with one's willingness to abide typos, but is not by any means the same. I personally was the loosest of all the "talkers" I knew, both in terms of leaving typos on the

screen and in terms of peppering my sentences with all sorts of silly abbreviations. For instance, I will now retype this very sentence as I would have in "talk mode," below.

F ins, I will now retype ts very sent as I wod hv in "talko mode," below.

Not bad! Only two typos. The point is, the communication rate is raised considerably—nearly to that of a telephone—if you type well and are willing to be informal in all these ways, but many people are surprisingly uptight about their unpolished written prose being on exhibit for others to see, even if it is going to vanish in mere seconds.

All of this I bring up not out of mere windbagery, but because it bears strongly on the Turing Test. Imagine the microscopic insights into personality that are afforded by watching someone—human or otherwise—typing away in "talk" mode! You can watch them dynamically making and unmaking various word choices, you can see interferences between one word and another causing typos, you can watch hesitations about whether or not to correct a typo, you can see when they are pausing to work out a thought before typing it, and on and on. If you are just a people watcher, you can merely observe informally. If you are a psychologist or fanatic, you can measure reaction times in thousandths of a second, and make large collections and catalog them. Such collections have really been made, by the way, and make for some of the most fascinating reading on the human mind that I know of. See, for instance, Donald Norman's article "Categorization of Action Slips" or Victoria Fromkin's book *Errors of Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*.

In any case, when you can watch someone's real-time behavior, a real live personality begins to appear on a screen very quickly. It is far different in feel from reading polished,

postedited linefuls such as I received from Nicolai. It seems to me that Alan Turing would have been most intrigued and pleased by this time-sensitive way of using his test, affording so many lovely windows onto the subconscious mind (or pseudomind) of the being (or pseudobeing) under examination.

As if it were not already clear enough, let me conclude by saying that I am an unabashed pusher of the validity of the Turing Test as a way of operationally defining what it would be for a machine to genuinely think. There are, of course, middle grounds between real thinking and being totally empty inside. Smaller mammals and, in general, smaller animals seem to have "less thought" going on inside their craniums than we have inside ours. Yet clearly animals have always done, and machines are now doing, things that seem to be best described using Dennett's "intentional stance." Donald Griffin, a conscious mammal, has written thoughtfully on these topics (see, for instance, his book *The Question of Animal Awareness*). John McCarthy has pointed out that even electric-blanket manufacturers use such phrases as "it thinks it is too hot" to explain how their products work. We live in an era when mental terms are being both validly extended and invalidly abused, and we are going to need to think hard about these matters, especially in face of the onslaught of advertising hype and journalese. Various modifications of the Turing Test idea will undoubtedly be suggested as computer mastery of human language increases, simply to serve as benchmarks for what programs can and cannot do. This is a fine idea, but it does not diminish the worth of the original Turing Test, whose primary purpose was to convert a philosophical question into an operational question, an aim that I believe it filled admirably.

Douglas R. Hofstadter is Professor of Cognitive Science at Indiana University in Bloomington. Hofstadter received the 1980 Pulitzer Prize for Godel, Escher, Bach: An Eternal Golden Braid (1979).