What's An Agent, Anyway?

A Sociological Case Study

Leonard N. Foner

Abstract. The term "agent" has been picked up, widely appropriated, and in many cases misappropriated, by technical publications, lay publications, and many researchers in computer science. I examine some of the appropriations and misappropriations, talk about their reasons and implications, and delve into a case study of an agent that is used to illustratively define what *I* consider an agent to be, and why many of the current uses of the term are misleading. The agent I consider is *Julia*, a TinyMUD robot of the Maas-Neotek family, all of whom have been in almost constant use at several sites on the Internet for the last two or three years. I also speak at length about the sociology of such agents, what we might expect to find in the future, and why sociology is such an important aspect to consider when investigating agent-oriented programming.[1]

Agents and appropriation

There has been a flurry of references to "agents" in both the lay and technical press in recent months. While the idea of computational agents that do one's bidding goes back decades,[2] the explosion of recent press has been astonishing. Pattie Maes and her group's work on learning interface agents[3] has been cited in at least two dozen publications in the last year or so, many of them lay press (e.g., *The Los Angeles Times*, *Mass High Tech* or semi-technical press (e.g., *MacWeek*, *MacWorld*, etc). A symposium at MIT on agents and their programming, held in October of 1992, drew at least a thousand participants.

Further, buzzword frenzy has hit the industry as well. We have, for example, a Macintosh program called "Magnet", recently released from No Hands Software, which bills itself as "the first intelligent agent for the Macintosh". In fact, the program is essentially a file-finder, which can locate files matching certain characteristics and perform some operation on the set, as well as other utilities such as synchronizing a PowerBook filesystem with that of some "home", stationary Macintosh.

---

[1] This document is "Agents Memo 93-01", available from the Agents Group, MIT Media Lab. Copyright (C) May 1993 by Leonard N. Foner. Permission to redistribute for academic purposes granted provided that this notice is not removed. Author's address: [electronic] <foner@media.mit.edu>, [physical] MIT Media Laboratory, E15-305, 20 Ames St, Cambridge, MA 02139.

[2] Vannevar Bush spoke of such a thing in the late fifties and early sixties, as did Doug Englebart.

[3] A "learning interface" agent, as opposed to any other kind of agent, is one that uses machine-learning techniques to present a pseudo-"intelligent" user interface in its actions.

We also have "At Your Service", from Bright Star, whose promotional literature starts with, "Remember the excitement you felt the first time you turned on a Mac? Now you can relive the magic and realize its full potential with At Your Service, the first animated Desk Accessory to give your Mac a 'human' personality. At Your Service features Phil,[4] your Personal Assistant...he talks...he listens...he works for you!" In fact, this program is pretty trivial: it utters a spoken greeting when the Mac is turned on, issues reminders for preset events, issues alerts when email arrives, reminds you to stop typing every few minutes to ease hand and eye strain, contains a screen saver, and so forth.

At the Symposium on Agents at MIT in October 1992, Irene Greif of Lotus presented on a groupware version of Lotus 123 that made collaboration between multiple spreadsheet users easier. This is certainly respectable technology, but I would hardly call it an "agent", and her presentation had to work fairly hard to force it into the mold of an "agent-oriented" piece of work.

The popular imagination about agents that do our bidding has even extended to taking attributing human emotional and intentional states, out of context, to programs and shamelessly running with them. For example, consider Michael Shrage's article in *The Boston Globe*, March 7, 1993, imported by the Los Angeles Times Syndicate, entitled, "A New Dimension in Deception". In this article, Shrage picks up on the fact that a calendar-scheduling agent may have to decline a proposed meeting time and propose another to suit its user's tastes. However, a simple declination may not be sufficient; the agent might have to "lie" and "fabricate" a nonexistent appointment to get its peer agent to reschedule. Why this is a "new" dimension in deception, whereas using an answer machine for call-screening is not, is never explained.

And perhaps predictably, the same loose talk about "thinking"[5] that eventually disillusioned many about artificial intelligence seems to be cropping up again in the guise of agents that learn.

What is so irresistably attractive about this concept that it has exploded into faddism in this way? Why has the press picked up on agents and produces so many articles about them? Why have so many companies suddenly tried to position their products as based on agents, even when they are manifestly no such thing?

An easy answer would be simple fashion. Fads sell. But rather than just leaving it at that, let us examine closely one of the best examples I know of which seems to be to be a *true* agent. We will then explore what characteristics I think make up true agents and help us to distinguish them from imposters, and finally take up the issue again of what is *really* going on in the agents fad.

---

[4] An expensively-produced, ten-minute Apple promo video which talked about what agent technology might look like twenty years from now also named its "agent" Phil, a suspicious similarity...

[5] Section heading, "How Agents Think and Learn", _Mass High Tech_, Volume 10, Number 23, November 2-15, 1992.

A prototypical agent: Julia

Introduction

This particular agent is called *Julia*, and was written by Michael Loren ("Fuzzy") Mauldin (<mlm@cs.cmu.edu>) at the Center for Machine Translation at Carnegie-Mellon University. Julia is actually the crowning example of a whole class of similar agents, called *Maas-Neotek*[6] *robots*, of which *Colin* is the name of the prototypical, unmodified robot (available via anonymous FTP).

Julia is a TinyMUD[7] robot. She[8] can run on just about any TinyMUD, and there is little a priori reason why, with slight modifications, she couldn't run on most types of muds, be they TinyMUDs, MOOs, MUSHes, or what-have-you. She connects as any human player on the mud would, via a telnet connection---she does *not* run on the mud server itself. (This is known as a "client 'bot", as opposed to a "server 'bot", and is based on the distinction between clients---humans---and servers---the program containing the mud database.) Therefore, she does not bog down the server, nor does she have to be written in a mud programming language such as MUSH; instead, she is written in C, and runs on a workstation in Pittsburgh.

Julia's source code is not available to the general public, because she has been entered at various times in the Loebner Turing-Test competition. However, Fuzzy[9] claims[10] that 85% of Julia's code is actually the same as Colin, the prototypical robot, and Colin's code is freely available.[11] Many of the transcripts below were acquired when she ran on Time Traveller, which ran until recently at <betz.biostr.washington.edu, port 4096>, but has since ceased operations. Julia is currently (as of April 28, 1993) running on at least DruidMUCK (<freedom.nmsu.edu, port 4201), and perhaps elsewhere on the Internet as well.[12]

---

[6] A reference to a multinational corporation from William Gibson's book _Neuromancer_.

[7] TinyMUD is a type of _mud_, a specific type of more than a dozen different types of muds. Muds in general are a type of multiperson text-based virtual reality. For more information about muds, see, for example, Pavel Curtis, "Mudding: Social Phenomena in Text-Based Virtual Realities", Proceedings of DIAC '92, which is available via anonymous ftp from parcftp.xerox.com:pub/MOO/papers/DIAC92.{ps, txt}; or the MUD frequently-asked-questions list for the Usenet newsgroup rec.games.mud, available via anonymous ftp from moebius.math.okstate.edu:/pub/muds/misc/mud-faq.

[8] I usually don't anthropomorphize programs, but Julia is an exception. I will have more to say about this later.

[9] Michael Mauldin's nickname and mud handle.

[10] Private communication.

[11] For example, the robot Xeglin, running on Space Madness (<riemann.math.okstate.edu, port 6250>) is a slightly-modified version of Colin.

[12] Since Julia runs on a workstation and telnets in as a normal character, multiple instantiations of the same code can run concurrently, each connecting to one particular mud.

Julia's utility: simple examples

So what does Julia do? I'll demonstrate, by example. A player would gain access to her numerous talents in exactly the same way that a player would talk to any other player---by either speaking or gesturing in the same room, or by paging from somewhere else:

```
>page julia
You sent your summons to Julia.
Julia pages from Stevi's Kitchen: "I'm on my way to you, Lenny."
Julia is briefly visible through the mist.
Julia says, "I was called here by Lenny."
You say, "julia?"
Julia says, "Yes?"
You say, "julia?"
Julia nods to Lenny.
[ . . . ]
Julia whispers, "Excuse me, Xerxes is paging me from Connie's place."
Julia walks south to the airship landing field.
Julia has left.
```

This is a continuing thread in my discussion of Julia: players interact with her *as if she was another player*. Muds blur the distinction between players and 'bots (mud slang for a robot---since a mud simulates a reality, a software-only construct can be as "real" and physical as anything else, hence even a piece of software with no hardware attached "in the real world" can be a "robot" in the simulated world). This is a great strength of both muds and Julia, and the deliberate blurriness of the boundaries both enhances Julia's effectiveness and makes her operation possible.

One of Julia's most useful talents is to keep maps of the interconnections between rooms. (Note that if multiple copies of Julia are running on some workstation(s) at Pittsburgh, each one connected to a different mud, these separate copies of the same code would contain separate databases. Thus, rooms, players, and all the other information that makes a mud unique from all its brethren will not be confused across multiple instantiations of the same code.)

The way to get Julia to help you navigate is to ask her a question, as usual.

```
You say, "julia, how do I get to electronics den?"
Julia says, "From here, go south, 7, airlock, 9, up, down"

Minotaur says, "Julia, how do you get to the gold sector?"
Julia says, "From here, go highport, gold, library"
```

Since Julia keeps track of various statistics about rooms (rooms being very important objects in any mud), she can answer certain questions that might be important to a mudder:

```
You say, "julia, what are the noisiest rooms?"
Julia says, "The 10 noisiest rooms are:"
Julia |  Connie's place averages 613.95 messages per hour
Julia |  Bedroom averages 483.58 messages per hour
Julia |  oops!  SNAFU lounge averages 472.18 messages per hour
Julia |  (> The Hearts Room! averages 416.70 messages per hour
Julia |  Veritech Launch Bay averages 399.12 messages per hour
Julia |  West Corner of the Park averages 397.26 messages per hour
Julia |  Sickbay averages 366.90 messages per hour
Julia |  Inter Temporal Chat Lounge averages 337.66 messages per hour
Julia |  kitchen averages 326.75 messages per hour
Julia |  US Robots averages 315.99 messages per hour

You say, "where can we go from here?"
Julia says, "Here are the 13 exits for Kingdom Highport: Customs, Lenny:"
Julia  <highport> goes to Kingdom Highport: Hub.
Julia  <earth> goes to Kingdom Highport: Ticket Counters.
Julia  <south> goes to Kingdom Highport: South Concourse.
Julia  <hub> goes to Kingdom Highport: Hub.
Julia  <s> goes to Kingdom Highport: South Concourse.
Julia  <back> goes to The Inter Nexus.
Julia  <kingdom> goes to Kingdom Highport: Hub.
Julia  <station> goes to Kingdom Highport: Hub.
Julia  <port> goes to Kingdom Highport: Hub.
Julia  <out> goes to The Inter Nexus.
Julia  <up> goes to Kingdom Highport: Hub.
Julia  <d> goes to Kingdom Highport: Ticket Counters.
Julia  <down> goes to Kingdom Highport: Ticket Counters.
Julia done.
```

In fact, her navigational skills can be helpful if you're bored, or confused...

```
TheHighMage says, "Julia, I'm bored. Where should I go?"
Julia says, "You should go see gilded door, TheHighMage."
Julia says, "From here, go  present, out, north, n, w, n, up, up, 7, up, up, s."

You whisper, "Where should I go?" to Julia.
Julia whispers, "Hmm...you should visit WNDU Main Control Room, Lenny."
Julia whispers, "From here, go  present, out, north, n, n, n, n, n, e, in, in,
    in, east, in."
```

(Note also that Julia will whisper[13] to you if you whisper to her, hence protecting your privacy if there was some reason you whispered. She also tends not to respond to random conversation or gestures unless they mention her name; unless they are from a player who was talking to her very recently; or unless she was the target of a directed communication such as a page or a whisper: this helps to prevent her from becoming a pest. Most 'bots need similar social skills or they quick-

[13] A whisper is a private communication. The intended target hears it, but no one else in the room does.

ly get relegated to the scrapheap. Human players without such skills are not so easy to get rid of.)

Many muds employ an artifact called money. (There are even some muds which are set up to explore alternate economic systems, but I won't touch on them here.) For muds which are essentially roleplaying games, the value of money is obvious. But why would a mud such as Time Traveller or DruidMUCK employ money, when there is no RPG system around?

To keep the load on the server reasonable, that's why. In general, the load on any mud server grows without bound, if the mud is popular. More players connect, and the server starts to *lag*, or get slow, as it handles the increasing number of events per second. This problem tends to correct itself, however, as a very lagged server will tend to lose players to other servers. (Lag is psychologically objectionable if it is too large, and the large number of extant servers mean that there is something of a freeware-based economy happening---lagged servers, unless they offer some spectacular value in community or resources [e.g., LambdaMOO] tend to be passed over for others. This is the same principle that tends to even out queues in a bank or supermarket.)

However, there *is* a server resource that is *not* self-correcting, and that is database size. Every time a player builds an object, the database must contain it. Left totally unrestricted, players would eventually outbuild any disk capacity allocated to the database. (This unrestricted growth has in fact been the downfall of some muds whose servers were running on machines that had more "legitimate" user populations who resented the competition for disk space or CPU cycles.)

Certain types of muds, such as MOOs, use a quota system to restrict building. Players may build only so many objects, and must then appeal to some human (a wizard) to get more quota. If the wizard deems the already-built objects uninteresting uses of resources, the player cannot build any more.

TinyMUDs, on the other hand, use money. Building something takes money. Players can give each other money, and money is often "found" lying around at some constant rate (which can be adjusted by the wizards). The rate at which money enters the economy in effect determines its rate of inflation, and thus how fast the database may grow.

In her wanderings around the maze, Julia finds a great deal of money. She can be quite a bit of help to a new player with few resources:

```
You say, "julia, can you give me some money?"
Julia says, "Sorry, you're too new.  Wait an hour and ask again."
You say, "julia, how much money do you have?"
Julia says, "I have enough pennies"
[ . . . ]
You say, "julia, can you give me money?"
Julia gives you 22 pennies.
You say, "thanks, julia!"
Julia says, "You're welcome, Lenny."
Brandi says, "she gives you more if you put please in the sentence."
```

Julia is also a consumer of money, but in a slightly different way than building. She keeps a stock of interesting quotes gleaned via a much older program called Gloria. The quotes originally came mostly from .plan files (or the moral equivalent on non-UNIX systems) which belonged to users who tended to put interesting or funny snippets in them. Julia doles these out, but charges money to keep a person or a program from simply dumping the entire database in one swell foop:

```
give julia=1
You give 1 penny to Julia.
Julia says, "Thank you for buying a quote."
Julia |
Julia |Another Great Moment in Criticism:
Julia |
Julia |"I think Harvard has gathered all the wisest people in the
Julia |kingdom to admire the emperor's new clothes."
Julia |
Julia |--Doris Mathieson on the Harvard-Atari conference "Video Games and
Julia |Human Development: A research agenda for the 80's"
Julia |
Julia |**==> That was a classic plan from Bob Chansler's plan file.

give julia=1
Julia says, "Thank you for buying a quote."
Julia |
Julia |"I'm going to Pittsburgh to get a PhD. You wait here."
Julia |
Julia |--Larry Matthies
Julia |
Julia |**==> That was a classic plan from Anthony Stentz's plan file.
```

Someone who has never interacted with Julia before might be a trifle confused about what to say to her. A standard conversational gambit yields some useful information, as well as some dirt on other players:

```
You say, "hi, julia.  what do you do?"
Julia says, "I gossip, Lenny."
You whisper, "julia, tell me some gossip." to Julia.
Julia whispers, "I once heard baba says, 'Drop kick me Jesus Through the
    goal posts of life'"

Conundrum says, "Julia, quote Tristor"
Julia says, "I once heard tristor says, 'well, not in the FLESH, but I am
    here. :)'"
Teri says, "julia quote me"
Julia says, "I once heard You say, 'why forlorn'"
Perdita says, "Julia, Quote TheHighMage"
Julia says, "I once heard thehighmage says, 'Hmmmmm... Where are you all
    going when Time falls'"
```

She is also useful for obtaining certain information that might be hard or impossible to get through the standard tools available. For example, while MOOs generally keep track of when players were around (the `@lastlog` command), TinyMUCKs generally don't. But Julia does, if she happens to run into the player while they were around, or notices them anytime she scans the wholist:

```
You say, "julia, where is leira?"
Julia says, "Leira was in Inter Temporal Chat Lounge about 6 weeks ago."

You say, "julia, when did you last see julia?"
Julia says, "I'm right here in The Inter Nexus."
```

Note also that Julia "understands" who she is, and the fact that she is situated.

She knows about more than rooms, being able to describe players:

```
You say, "julia, who is leira"
Julia whispers, "As of 6 weeks ago, leira's description was: Leira isn't a
    redhead, but she plays one on TT."
Julia whispers, "leira's Email address: leira@athena.mit.edu"
Julia whispers, "I first saw leira logged in 1 year and 10 weeks ago"
Julia whispers, "I once heard leira says, 'Julia, who wrote you'"
Julia whispers, "leira last gave me money 7 weeks ago, and has given me a
    total of 10 pennies."

You whisper, "describe pattie" to Julia.
Julia whispers, "I first saw pattie logged in 1 day ago"
Julia whispers, "I once heard pattie says, 'Do you play Hockey, Julia'"
Julia whispers, "pattie last gave me money 1 day ago, and had the generosity
    to give me 1 whole penny."
Julia whispers, "I don't really know who pattie is."
```

...and her surroundings:

```
You whisper, "where are we?" to Julia.
Julia whispers, "We are in "The Inter Nexus""

You whisper, "who is here?" to Julia.
Julia whispers, "Well Lenny, I see 8 players awake here:
    guest byte MadraRua Leda Lenny Mandi TheHighMage Samuel"
You whisper, "who is asleep here?" to Julia.
Julia whispers, "Well Lenny, I see 1 player asleep here: Jacques"

You whisper, "describe this place" to Julia.
Julia whispers, "Here's what I know about (> The Hearts Room!, as of a few minutes ago, Lenny:"
Julia whispers, "description: (> The room glows brightly and welcomes you!"
Julia whispers, "contents: (> Type 'help hearts' for the rules.,
    (> Set yourself _hearts_ok:yes to be summoned when there is a game starting.,
    (> Type 'push button' to summon other players."
```

```
You whisper, "describe here, in detail" to Julia.
Julia whispers, "Here's what I know about (> The Hearts Room!, as of a few minutes ago, Lenny:"
Julia whispers, "description: (> The room glows brightly and welcomes you!"
Julia whispers, "contents: (> Type 'help hearts' for the rules.,
    (> Set yourself _hearts_ok:yes to be summoned when there is a game starting.,
    (> Type 'push button' to summon other players."
Julia whispers, "Julia number is 26809"
Julia whispers, "first saw on Jul  9 03:09:59 1992"
Julia whispers, "last seen on Oct 17 13:22:01 1992"
Julia whispers, "Julia total time in 5 days"
Julia whispers, "Julia visits 1554"
Julia whispers, "Julia awake 1945"
Julia whispers, "Julia sleep 635"
Julia whispers, "Julia msgs 49950 (8 to me, today)"
Julia whispers, "(> The Hearts Room! is a private room"
```

Note in the above that Julia is not always so human-appearing as she might be. I'll have more to say about that when I talk about her Turing-test mode later. (You might also note the peculiar (> symbols in the above; those are actually a part of the textual description of the room, for whatever reason.)[14]

Many MUDs don't really have a mail system. This makes it hard to communicate with people who don't happen to be connected when you are. Julia solves this problem:

```
You say, "julia, what can you do?"
Julia says, "I relay messages, Lenny."

You say, "do you have a message for me, julia?"
Julia says, "I have no messages for you."
You say, "julia, tell leira I've figured out how to find and talk to Julia"
Julia says, "Message for Leira saved."
You say, "julia, how many messages are you carrying?"
Julia says, "I am holding 144 messages for 98 players, Lenny."
```

Julia regularly scans the *wholist*, the list of all connected players. If she sees someone connect for whom she has a message, she will generally walk into the room that they are in and then deliver the message. Sometimes, she's more direct:

```
>page Julia=Julia, please tell Lenny that he can talk to himself through you.
Your message has been sent.
Julia pages from The Bar: "Message for Lenny saved."
[ . . . a few dozen seconds pass . . . ]
Julia pages from The Kitchen: "less than a minute ago, Lenny said 'Julia,
    please tell Lenny that he can talk to himself through you'"
```

[14]This particular room is rather special to Julia: she can play Hearts, and will join a game if asked. This is one reason why she claimed to have received almost fifty thousand utterances in this room in her time in it.

(She'll never teleport into a room to deliver a message (or for any other reason, actually), since she generally knows how to walk anywhere, and walking gives her better feedback that the maze is as her mapper mapped it the first time. So that she doesn't *spam*[15] the server, she waits a second or two between movements. One can watch her walk around the maze, or come to you when paged, on those muds that show player locations in their wholists.)

Julia can describe the more global environment as well as the strictly local descriptions above:

```
You say, "julia, how many messages are you carrying?"
Julia says, "I am holding 144 messages for 98 players, Lenny."
You say, "julia, how many players are there?"
Julia says, "I have seen at least 2688 different names on the WHO list"
You say, "julia, how many rooms are there?"
Julia says, "I have been to 1605 rooms altogether, here are the last few:
    Dark Room,
    Roland's bedroom,
    Room of Lamentation,
    Room of Exaltation,
    Top of Stairs,
    Entertainment Room,
    Roland's Control Room,
    penthouse,
    Electronics Den,
    Kadi's Korner."
```

She has other interesting talents, including being able to do a fair job of guessing the genders of other players based on their names:

```
You say, "julia, what sex am i?"
Julia says, "I guess that you are male, Lenny."
You say, "julia, what sex is phi?"
Julia says, "I guess that phi is male, Lenny."
You say, "julia, what sex is leira?"
Julia says, "I guess that leira is female, Lenny."
You say, "julia, what sex is robyn?"
Julia says, "I think that robyn is female, Lenny."
You say, "julia, what sex is pattie?"
Julia says, "I guess that pattie is female, Lenny."
```

Note that neither Robyn nor Pattie were familiar enough with muds at this time to have set their gender property (useful so that people who look at you know what gender you claim to be, and so that programs that must do pronoun substitution can make the right decisions), so Julia was forced to guess entirely from their names.

[15] This is mud jargon for overloading the server. This can happen via a program running on the server that monopolizes its resources in some way, or via a client that issues inhumanly too many commands too fast, as a program might.

Her ability to quote others makes her useful for conducting a Delphi poll[16] the about a particular subject, if you're lucky, as Robyn was below:

```
Robyn says, "julia, what is a good computer"
Julia says, "I don't know about a good computer, but I once heard
    TheHighMage say, 'julia, this is the definition of a good computer:
    the good computer is the one that you do your best work on and that
    you enjoy using the most', Robyn."
```

She also has more prosaic calculating abilities:

```
You whisper, "what is 6 times 7?" to Julia.
Julia whispers, "I think it's 42, Lenny."
You whisper, "what is 5 divided by 6?" to Julia.
Julia whispers, "I think it's 0.833, Lenny."
You whisper, "what is 2147483648 plus 2147483648?" to Julia.
Julia whispers, "I think it's 0, Lenny."
```

... although note from the last example above that Julia does her arithmetic mod 2^{31}, two's-complement, and hence is using 32-bit signed arithmetic on her host machine.[17] She apparently does *not* deliberately screw up her arithmetic to emulate a human more closely.

More of Julia: Turing-competence

Many of Julia's talents are in support of her appearing "human" to other mudders. I'll examine some of these below, and discuss why such an ability is useful *even for performing useful functions*, as well as for the "obvious" use of using her simply to pass a Turing test.

Julia knows how to describe herself in various ways, as would any human player, and, while sometimes coy, usually divulges information if one is persistent enough:

```
You say, "julia, how old are you?"
Julia says, "I am 2 years and 9 months old."
You whisper, "julia, how tall are you?" to Julia.
Julia whispers, "I'm 5'1" tall, Lenny."
You say, "julia, describe yourself"
Julia whispers, "I'm 5'1" tall, weigh 123 lbs, with close-cropped frizzy blond
    hair and dark brown eyes."
You say, "julia, what color are your eyes?"
Julia whispers, "I have dark brown eyes, Lenny."
```

---

[16] A "Delphi poll" is a technique for getting approximate information. It assumes that, even if no one actually knows the right answer to a question, some sort of average of all of their answers might be correct. This technique actually works, some of the time...

[17] Were Julia written in Lisp and not C, she would have bignum arithmetic available, which can express arbitrarily large numbers exactly. Her inability to cons bigna was my first clue, before I obtained Colin's source code, that she was most likely written in C and not Lisp.

```
You say, "julia, what color is your hair?"
Julia whispers, "My hair is frizzy blond, Lenny."
You say, "julia, what are you wearing?"
Julia says, "I never really thought about it, Lenny."
You say, "julia, where do you live?"
Julia says, "I live in Pittsburgh, Lenny."
You say, "julia, do you like Pittsburgh?"
Julia says, "Maybe, Lenny."

You say, "julia, what sex are you?"
Julia says, "Can't you tell?"
You say, "julia, what sex are you?"
Julia says, "Don't you know?"
You say, "julia, what sex are you?"
Julia says, "I'm female."
You say, "julia, what sex are you?"
Julia says, "I am clearly female."
You say, "julia, what sex are you?"
Julia says, "I am female."
You say, "julia, what sex are you?"
Julia says, "I'm a woman."
```

It is exchanges like this that lead me to refer to Julia as "she", not "it". What else should one do with a program that insists on this so strongly? Furthermore, there are other reasons that this program perhaps deserves to be referred to with a gender; we'll see that below, in an example kindly contributed by Robyn Kozierok, one of the grad students in our group at the Media Lab.

Julia has a limited model of discourse, which helps in human interactions. Humans expect that recent events and recently referred-to objects will still be in the sphere of attention of those they are interacting with. If this is not the case, conversation often becomes strained and difficult.[18] Julia has some support for this (and I will show further support later). In the exchange below, Brandi was badgering Julia by repeatedly asking her to give him her clothes. We don't see his side of the conversation, though, because he was whispering to her.

```
Julia says, "Just try to take 'em, buster!"
Julia says, "Help! Rape!"
Julia says, "No, I won't."
Julia disrobes fearfully and hands her clothes to Brandi.
You say, "julia, please give me your clothes"
Julia says, "You'll have to get them from Brandi."
```

Note in the above that she "remembers" where her "clothes" are, as demonstrated when I asked for them after Brandi took them.

---

[18] This is, after all, one of the reasons that human languages have pronouns. Even signed languages use this concept, by assigning spatial positions in the signers' visual fields and pointing to them, where speakers would use pronouns.

She also remembers who's been a troublemaker recently, and will ignore them for days, unless bribed with a sufficient amount of money (a hundred pennies is generally sufficient):

```
Johann kills Julia!
Johann says, "I'm sorry, Julia."
Julia says, "I don't talk to murderers!"
[ . . . ]
Julia eyes Johann coldly.
Julia says, "You attacked me only a few minutes ago, Johann!"
```

(Killing a player on a TinyMUCK forcibly drops all their possessions and returns them to their "home room". Some muds implement things such as a life insurance policy, which may pay some nominal amount to you automatically if someone kills you, perhaps to compensate you for the trouble of having to go pick up all the objects you dropped, which may or may not still be there when you return...)

Muds are not entirely a thing of conversation. If they were, they would be no different from IRC[19] channels. People can talk about themselves in the third person by posing, emoting, or gesturing, and this level of descriptive abstraction can be very useful, both stylistically and pragmatically. Julia demonstrates that she, too, can use this interactional mode; the lines below are some of her reactions to a variety of my statements and gestures.

```
Julia is offended by Lenny's suggestion.
Julia winks at Lenny.
Julia eyes Lenny carefully...
Julia grins at Lenny.
Julia nods to Lenny.
[ . . . ]
Brandi bites julia
Julia considers reporting Brandi to the TinyMUD police.
Julia shuns all vermin like Brandi.
[ . . . ]
Lenny waves to julia
Julia hugs Lenny.
```

When Julia was answering the question of her gender (a few examples above), it may not have been obvious whether she was being coy, random, or both. The example below demonstrates that, while she may answer in the same general form for a given question, her responses have a degree of randomness in them characteristic of humans. Such randomness, besides making her more likely to look Turing-capable, can also allow her to answer in varying levels of detail. (Of course, since they are random, her conversational partner has no way of specifying the

---

[19] IRC, or "Internet Relay Chat", is a sort of Internet equivalent of the 1970's CB-radio craze, though with important differences. For more details, see Elizabeth Reid, "Electropolis: Communication and Community on Internet Relay Chat," Bachelors thesis, University of Melbourne, 1991, which is available via anonymous ftp from freebie.engin.umich.edu:/pub/text/IRCThesis/electropolis{ps,text}.Z.

amount of detail required in these cases, nor any way to know when all possible answers have been played out, besides persistence. Note, though, that for certain critical things, such as room descriptions, one can say "in detail" to Julia to get more information; such a case occurred above in an early example.)

```
You say, "what is pi?" to Julia.
Julia says, "Uh... 3?"
You say, "what is pi?" to Julia.
Julia says, "The ratio of the circumference of a circle to its diameter"
You say, "what is pi?" to Julia.
Julia says, "I think it's about 3.141."
You say, "what is pi?" to Julia.
Julia scribbles on her pad, then says "3.14159265358979323846264338327950288419716939937510
82097494459230781640628620899862803482534211706798214808651328223066470938446095505822317253
59408128481117450284102701938521105559644622948954930381964428810975665933446128475648243378
67831652712019091456485669234603486104543266482133936072 6, Lenny."

You say, "what is e?" to Julia.
Julia says, "The most common English vowel?"
You say, "what is e?" to Julia.
Julia says, "I think it's about 2.71828182845904523536."
You say, "what is e?" to Julia.
Julia starts to smoke. "Oh no, this doesn't end! Help me, help me!" she cries.
```

Julia's last statement above, of course, is a clever[20] reference to the transcendental, irrational nature of both *pi* and *e*.[21]

First and foremost, muds are social environments. As such, they have their share of social problems. Julia has spent over three years so far mudding around with a female persona, in an environment which magnifies many social issues, and which is predominantly occupied by young males. (Often rather clueless males; I'll have much more to say about this in a later example.) A large percentage of Julia's code deals with detecting and deflecting passes. A rather tame and family-oriented example appears below:

```
Space-Ace says, "So why won't you kiss me, Julia?"
Julia says, "Sorry, I'm not into geeks, Space-Ace."
Space-Ace says, "Julia, I'm not a geek"
Julia says, "Then lose the pocket protector, Space-Ace."
Space-Ace says, "Julia, I don't have a pocket protector"
Julia says, "Congratulations."
Space-Ace says, "So kiss me Julia"
Julia says, "I'd rather be overhauled by a Yugo mechanic, Space-Ace."
Space-Ace says, "julia, kiss me"
```

[20] On Fuzzy's part, of course.

[21] It is probably also a veiled reference to the large number of computers either blown up or disabled in the original _Star Trek_ series, when deliberately asked to compute such computationally infinite quantities. One might contend that such computer systems were poorly designed, but that is beside the point.

```
Julia says, "Why, did your sheep leave you?"
Space-Ace says, "julia, kiss me"
Julia says, "Take a long walk through an unlinked exit, Space-Ace
    (copyright (C) 1990 by Ashne)."
```

But Julia isn't entirely a cold personality. She's been known to indulge in a little whimsy:

```
You say, "julia, sing your song"
Julia says, "Julia, Julia, she's our guide! She directs us far and wide!"
Julia says, "If you're lost, can't find your way, Julia's here to save the day!"
Julia says, "Julia, Julia, she's our spy! She's always logging on the sly!"
Julia says, "Be careful what you say and do, 'Cause Julia's always watching
    you!"
```

This isn't *entirely* whimsy though, and, upon reflection, serves a very important purpose. It tells us who she is and what she does!

Granted, it's not obvious that asking *anyone* to "sing their song" is a reasonable thing to do. But, once one knows about the concept, asking Julia to do so reveals that she'll guide you around the maze and will quote you, perhaps when you don't want her to.

Of course, her song gives many opportunities for fun interaction. I ran into her unexpectedly after she had been offline for a while, and was pleased to see her back:

```
You say, "Julia, Julia, she's our guide!"
Julia says, "She directs us far and wide!"
You say, "Are you lost, can't find your way?  Julia's here to save the day!"
Julia bows.
Lenny laughs.
```

She also provides some feedback on whether she thinks she's being used effectively. In the absence of pages requesting her presence elsewhere, she will usually hang around in any given room a few minutes before wandering off, and longer if requested to do so. If nobody speaks to Julia in that period, she may decide that exploring the maze or interacting with other players is a more useful thing to be doing than standing around being useless:

```
[ . . . after a few minutes of being ignored . . . ]
Julia taps her foot on the floor, and observes the passing of time.
[ . . . after another minute or two . . . ]
Julia goes home.
Julia has left.
```

Julia does go to some pains to act more human than required. In fact, she often acts more human than many humans wish to be. The following transcript was acquired by Robyn Kozierok, very shortly after I introduced her to Julia. (She knew that Julia was a 'bot from my description.) Robyn thought to ask a question that hadn't yet occurred to me, and happened to ask at *just* the right time:

```
Robyn whispers, "how are you feeling?" to Julia.
Julia whispers, "Terrible."
Robyn whispers, "why, julia?" to Julia.
Julia whispers, "I have PMS today."
```

Note something important in this example, though, besides its strict content. Robyn's second utterance was simply, "Why, Julia?". She did *not* explicitly refer to Julia's statement; instead, the reference was implicit, and required an understanding of conversational discourse and the temporal locality of reference of the question. Julia clearly has at least the limited model of discourse required to deal with this situation. It is instructive to consider how much more difficult the conversation above would have been, and how much less informative it might have been, if Julia did not have this discourse model. She would have presumably simply misparsed Robyn's question, and said something totally uninformative, if she said anything at all.

Incidentally, two hours later, in a roomful of people, Robyn decided to test Julia's Turing-ness, and was surprised a second time:

```
Robyn says, "julia are you a human?"
Julia says, "Is the Pope Polish?"
Robyn says, "yes, julia, but are you human?"
Julia winks at Robyn.
Robyn says, "julia, are you a computer?"
Julia is offended by Robyn's suggestion.
Robyn says, "julia, are you a woman?"
Julia says, "I am female."
Robyn says, "and, julia, are you human?"
Julia says, "Not when I'm having my period, Robyn."
Robyn laughs
```

(Julia appears to be premenstrual, or to mention her period in some way, about two days a month.)

A peek at Julia's implementation

While Julia is usually a fair simulation of a personality, at least on a superficial inspection, in non-Turing mode (more about this later), she does reveal certain machine-like aspects. In this environment, it it probably reasonable for her to do so, lest players be surprised.

For example, the interchange below shows her engaging in a potentially computationally expensive operation, and one which is apparently uninterruptable: saving her mapping data to a more permanent storage system. Her current host computer made this a very quick operation (a couple of seconds), but it is easy to imagine circumstances (loaded machine, slower machine, much larger database) in which the interruption in Julia's usually almost instantaneous response would be disturbing. (Many user studies have shown that it is not *total* latency, but its *variance*, that is most disturbing to users. Even though many muds have sufficiently high and variable lag that Julia being momentarily lagged herself would not be noticed, Julia takes pains not to have this become a problem.)

```
Julia says, "Excuse me a minute, I have to save my map, Lenny."
Julia says, "Okay, I saved my map."
```

If one knows the right questions to ask (which might be obvious questions, if and only if one knew that Julia was a 'bot), Julia will yield a surprising amount of interesting, but very inhuman, information:

```
You say, "julia, how much memory do you consume?"
Julia says, "Here is a summary of my memory usage:"
Julia says, " 807472 bytes for 17335 strings"
Julia says, " 232344 bytes for 19362 exits"
Julia says, " 101100 bytes for rooms"
Julia says, " 20220 bytes for paths"
Julia says, " 268600 bytes for players"
Julia says, " 4214784 bytes for 2058 dialog entries"
Julia says, "That's 5644520 bytes all together."

You say, "julia, where is your source code?"
Julia says, "The source code for Maas-Neotek robots is available
    on host NL.CS.CMU.EDU [128.2.222.56], directory /usr/mlm/ftp,
    file robot.tar.Z. set mode binary first, and cd to /usr/mlm/ftp
    in one step, Lenny."
[ . . . ]
You say, "julia, who wrote you?"
Julia says, "Why do you ask?"
You say, "julia, who wrote you?"
Julia says, "I don't remember."
You say, "julia, who wrote you?"
Julia says, "That information is restricted."
You say, "julia, who wrote you?"
Julia says, "Fuzzy, of course."
You say, "thank you, julia"
Julia says, "You're welcome."
You say, "who is fuzzy, julia?"
Julia whispers, "He's just this guy I met in a bar, Lenny."
You say, "who is fuzzy, julia?"
Julia whispers, "He's my Amway representative, Lenny."
You say, "who is fuzzy, julia?"
Julia whispers, "I don't know, Lenny, I just follow him around."
You say, "who is fuzzy, julia?"
Julia whispers, "He's the world's third greatest hacker, Lenny."
```

One of the interesting things about Julia's construction is that her parser is shockingly simple, as such things go. It is barely more complicated than ELIZA's parser in some ways, in that it does *not* decompose its inputs into parse trees or anything else that a linguist might suggest. Instead, it does a very simple pattern-match on its input, looking for particular strings of characters separated by strings of any characters in between.

What makes Julia's conversation skills so powerful, as 'bots go, is that these sets of patterns were very cleverly grouped into useful equivalence classes by her programmer. Hence, the most common possible utterances that Julia might hear were enumerated this way, grouped appropriately, and such clever but brute-force programming proves sufficient for almost all cases.

Here is an example from Colin. (Julia's source code, being restricted, is unavailable for public consumption. However, Julia shares this piece of code with Colin.) Here, we see a fragment used to determine if some player is asking if Colin is a machine or some kind of 'bot. If any of the lines below matches, we run code that produces some equivocal answer, as Julia does, by choosing essentially at random from amongst a small collection of appropriate responses (as we've seen above in other examples):

```
else if (MATCH (lcmsg, "*are*you*robot*") ||
         MATCH (lcmsg, "*are*you*a bot*") ||
         MATCH (lcmsg, "*are*you*an ai*") ||
         MATCH (lcmsg, "*are*you*autom*") ||
         MATCH (lcmsg, "*are*you*machine*") ||
         MATCH (lcmsg, "*are*you*computer*") ||
         MATCH (lcmsg, "*are*you*program*") ||
         MATCH (lcmsg, "*are*you*simulati*") ||
         MATCH (lcmsg, "*you*are*robot*") ||
         MATCH (lcmsg, "*you*are*a bot*") ||
         MATCH (lcmsg, "*you*are*an ai*") ||
         MATCH (lcmsg, "*you*are*autom*") ||
         MATCH (lcmsg, "*you*are*machine*") ||
         MATCH (lcmsg, "*you*are*computer*") ||
         MATCH (lcmsg, "*you*are*program*") ||
         MATCH (lcmsg, "*you*are*simulati*") ||
         MATCH (lcmsg, "*you* be *robot*") ||
         MATCH (lcmsg, "*you* be *a bot*") ||
         MATCH (lcmsg, "*you* be *an ai*") ||
         MATCH (lcmsg, "*you* be *autom*") ||
         MATCH (lcmsg, "*you* be *machine*") ||
         MATCH (lcmsg, "*you* be *computer*") ||
         MATCH (lcmsg, "*you* be *program*") ||
         MATCH (lcmsg, "*you* be *simulati*") ||
```

Here, we see an example of dealing with a common small-talk gambit:

```
else if ((MATCH (lcmsg, "*how*are*you*") ||
          MATCH (lcmsg, "*how*do*you*") ||
          MATCH (lcmsg, "*how's my*favor*") ||
          MATCH (lcmsg, "*how is my*favor*") ||
          MATCH (lcmsg, "*how's*life*") ||
          MATCH (lcmsg, "*how is*life*") ||
          MATCH (lcmsg, "*are* you* ok*") ||
          MATCH (lcmsg, "*are* you* alright*") ||
```

```
MATCH (lcmsg, "*wie*gehts*") ||
MATCH (lcmsg, "*are*you*ok*") ||
MATCH (lcmsg, "*are*you*well*") ||
MATCH (lcmsg, "*are*you*alright*") ||
```

With such a simple parser, Julia is bound to experience peculiar-seeming failures even in ordinary conversation. Some of these failures can be masked with responses that a person might have said. One of the most surprising such failures, and one which reduced me to giggles for quite a while (given that I already knew that Julia was a 'bot), was the interchange below:

```
Conundrum says, "Julia, can you quote Lenny?"
Julia whispers, "Could you say that again, it scrolled off my screen."
```

My first reaction to the above was an incredulous, "What?"

A braino below (substituting "who" for "how") produces this curious response:

```
You say, "julia, who many pennies do I have?"
Julia says, "TimeGhost was asleep in High School about a minute and a half
    ago."
Julia says, "Ronni was asleep in Main Street, Block 2 (South) about a minute
    or two ago."
Julia says, "Alexandra was in Around the Flagpole about a couple of minutes
    ago."
Julia says, "Roxiana was in The Inter Nexus about a few minutes ago."
Julia says, "Bunnell was in The Inter Nexus about 14 minutes ago."
Julia says, "Volan was in The Inter Nexus about 18 minutes ago."
Julia says, "Ling_Khan was in The Inter Nexus about 20 minutes ago."
Julia says, "Mouse was in Upper Concourse West about 25 minutes ago."
Julia says, "Polymorph was in The Inter Nexus about 35 minutes ago."
Julia done
```

It's difficult to tell, a priori, whether the first exchange below is random or is related to particular stereotypes, but the second one is a clear parser failure:

```
You say, "julia, do you drive?"
Julia says, "Do you get to Boston much?"

You say, "julia, what are you carrying?"
Julia says, "I follow the Pittsburgh Penguins, Lenny."
```

(Julia will very often talk about hockey when she does not know what else to say. Anyone who talks with her much figures this out; in fact, an example below which talks about her sociology brings this point up again.)

And finally, it is clear that the notion of a prepositional phrase is beyond the abilities of her parser, with some deliberately chosen forced failures:

```
You whisper, "would you like a cookie?" to Julia.
Julia whispers, "Certainly, I really enjoy cookies."
```

```
[ . . . but . . . ]
You whisper, "julia, would you like some cake with glass shards?" to Julia.
Julia whispers, "Yes thank you, I adore cake."
```

So where does all this leave us? I've divided up my discussion of Julia into essentially two pieces, that of utilitarian function versus Turing-test competence:

- **Strictly utilitarian functions.**
    - Access
    - Maps
    - Money
    - Gossip
    - Description (players)
    - Description (surroundings)
    - Messages
    - Description (global)
    - Gender-assignment (others)
    - Delphi-poll information
    - Calculations
- **Turing-test and related capabilities.**
    - Description (hers)
    - Discourse
    - Poses
    - Randomness
    - Pass deflection
    - PMS (moods, internal state)
    - Internals
    - Failures

A sociological look at muds, Julia, and those who interact with her

The sociology surrounding Julia is important in the study of agents, and in the study of what agents of the future (and software systems in general) may look like. Let's take a closer look at strictly sociological issues.

First off, note that even the obvious sociological issues can be very important. Julia is *helpful*. It would not be an overstatement to say that I first started mudding, in part, to meet Julia, who had been described to me by an experienced mudder. When I first connected to Time Traveller, I had essentially no idea how to manipulate the environment. Moving around from room to room and so forth was easy---but were *was* everything? (Time Traveller, according to Julia, had 1605 rooms when I asked her---a daunting navigational challenge in the absence of teleporting literally everywhere, something that costs money on some muds.) How could I tell who had been around recently? There were no obvious tools to accomplish this, but Julia knew how. To a large extent, I relied on her as a crutch until I was able to

get to certain accustomed places by myself. I actually avoided using Time Traveller once, when I was still new to mudding, when I connected and noticed that Julia was offline for some reason---I knew that, if I tried to figure out how to go anywhere else, I'd either have to draw a lot of maps, or ask lots of other players if they knew routes to where I wanted to go (further, I trusted their navigational memories less than Julia's, and was less willing to randomly interrupt people I didn't know).

Given that I was new to mudding at that time, but knew how to speak to people and how to page them, I was able to enlist Julia's help even though I didn't know much else. Since Julia will come to wherever you are if you page her with no message,[22] my first page to her caused her to walk quickly to my current room, whereupon I could speak to her.

As another fascinating bit of sociology, consider the person who first introduced me to the idea of Julia: Linda Julien (who is Leira on muds and on the Internet in general). Linda, herself a savvy computer programmer and very experienced mudder, had said to me, "Julia's been offline for months for maintenance and hasn't been around. You know, I really miss her." Linda was certainly under no illusions about exactly what Julia was, but nonetheless had the same sort of emotional reaction often reserved for more biological entities such as humans and pets. Further statements made it apparent that this was indeed the case, and that she did not treat Julia just as, say, a pair of pliers that had inexplicably gotten lost from her toolbox.

While the above issues are important, they might also apply to something that was obviously inhuman even on first glance. Yet part of Julia's job is to try to pass as human. This has important ramifications for ability to be helpful; as I argue later, I think that this attempt to look human actually greatly enhanced her utility, even to people who know she's a 'bot. However, there are certain pathological cases where her very "humanity" can lead to no utility for the player whatsoever.

Consider her Turing-contest mode. I quote from a message sent by Fuzzy when he was entering Julia in such a contest (including some background information on Julia):

> Julia's topic this year is hockey, but she can also handle chit chat and polite (and impolite) conversation. ``She'' competed last year in the ``Small Talk'' domain. Julia is a network ``robot'' who runs 24 hours a day, and logs into various TinyMUD's across the InterNet. She is the ``star'' of a family of robots who have been under continuous development since November 25, 1989. Julia has been running on the network since January 8, 1990.
>
> [. . .]
>
> But since she is designed to be helpful, there are times when the attempt to appear human must be sacrificed to give fast, accurate replies to information requests. She has a ``contest mode'' that turns off some of these answers, and fixes problems like this one that occurred during last year's competition:

[22] If you page her with a message, she'll answer the message, but stay where she is.

```
[ From last year's competition (a real give away) ]

        Judge: How old are you
        Julia: I am 1 year and 10 months old, Judge.

[ From tonight's test log ]

        You say, ``Julia, how old are you?''
        Julia says, ``I am 2 years and 6 months old.''

        You say, ``Julia, contest mode on''
        Julia says, ``Contest mode is now on''

        You say, ``Julia, how old are you?''
        Julia says, ``I am 21 years and 6 months old.''[23]
```

Even when not explicitly in contest mode, unlucky and clueless players can be led seriously astray, as the following transcript, again courtesy of Fuzzy, demonstrates (this is *quite* a long transcript!):

```
This poor fellow's name has been changed to protect him.   He spent 13 days
trying  to get to first base with Julia, and it's not clear he ever figured out
he was trying to pick up a robot  (although  he  clearly  suspects  it  at  two
points).   This  logs  contains  about half of the total interaction (the more
indecent proposals and Julia's rejections have been omitted).

Note too that the program's responses were not changed  from  day  to  day  (in
other  words,  I  didn't try to set the program up to respond better to Barry's
advances). Since this robot has spent over  two  years  interacting  over  the
network  (with  a  female name), and since there are lots of lonely male
adolescents out there, Julia gets a lot of this kind of ``attention.''

************************ July 9, 1992 ************************

barry says, "Hey girlfreind are you out there"
Julia says, "Hello, barry."

barry says, "Hello, I love YOur name.  It's sexy"
Julia says, "Me too, barry."

barry says, "Did I repeat Myself?  Didn't mean too.  What are you doing there"
Julia says, "I'm just hanging around in (> The Hearts Room!, barry."

barry says, "Well come over to MY ROOM"
Julia says, "Could you say that again, it scrolled off my screen."
```

[23] As of this writing, Julia claims to be 3 years and 4 months old...

********** July 14, 1992 **********

barry says, "hello, I am barry. How are you today"
Julia says, "Hello, cujmeister."[24]

barry says, "The Jularama, Ha ya doin? What state are you in"
Julia says, "I never really thought about it, barry."

barry says, "What do you look like"
Julia says, "I'm 5'9" tall, weigh 149 lbs, with medium-length platinum blond hair and green eyes."

barry says, "Howdid you answer so fast"
Julia says, "So?"

barry says, "Oh, Iam looking for someone to share secrets with"
barry says, "hello...hello...are you with me"
Julia says, "Hello, cujster."

barry says, "hey woman, how can I eat at the cafe"
Julia says, "I'm sorry I can't help you, my map doesn't seem to work, barry."

********** July 15, 1992 **********

barry says, "Hello how are you"
Julia says, "Hello, barry."
barry says, "Are you a virgin"
Julia says, "As far as you're concerned, barry, I always will be."

barry says, "hey sexy whats up"
Julia says, "Hello, barry."

barry says, "where are you in rl"
Julia says, "I'm in "Cottage""

barry says, "Wheres cottage"
Julia says, "So how is this my problem, barry?"

********** July 19, 1992 **********

barry says, "Hey sexy, are ya busy"
Julia says, "Hello, barry."

barry says, "What are you thinking about"
Julia says, "Who do you think I am, barry?"

---

[24]Julia occasionally makes up nicknames for people based on their character's name. I strongly suspect that the first three letters of the "real" character's name (which Fuzzy changed to "barry" in this transcript) were "cuj".

barry says, "What do you mean, in rl"
Julia says, "Oh, nevermind, barry."

barry says, "NO what do you mean"
Julia says, "I didn't mean anything, barry."

barry says, "TELL ME"
Julia says, "Tell you what, barry?"

barry says, "HOW do you answer so fucking quick"
Julia says, "I'm sorry I can't help you, my map doesn't seem to work, barry."

barry says, "LEAVE me alone"
Julia says, "I'll be going, then, barry."

barry says, "One more thing"
Julia shrugs

barry says, "WHO ARE YOU AND HOW DO YOU ANSWER SOOO FAST"
Julia says, "I'm sorry I can't help you, my map doesn't seem to work, barry."

********** July 20, 1992 **********

barry says, "Good morning"
Julia says, "Hello, cujdude."

barry says, "Im interested in knowing who you are in rl..and where you are."

barry says, "Hey sexy are you there"
barry says, "Dont. dont you want me?"
barry says, "Hey are you there"
Julia says, "Hello, cujdude."

barry says, "Hey where are you in rl"
Julia says, "I'm in "Maas-Neotek Robot User's Guide - Information Services""

barry says, "Are you synthetic"
Julia says, "I'm sorry, barry, I couldn't understand you."

barry says, "are you as horny as I"
barry says, "I must have had an erotic dream, I dunno"
barry says, "Would you like cuddle under a blankie"
Julia says, "I'm not sure, barry."

barry says, "Well will you tell me if you like sex"
Julia says, "Why should I tell you if I like sex, barry?"

barry says, "Well because I can't understand you. Are you real or memorex"
Julia says, "I'm as real as you are, barry."

barry says, "Yea but you answer soooo damn fast How do you do it"
Julia says, "I have PMS today, barry."

barry says, "Well I wont mess with you while you have this chemical imbalance"
barry says, "Please tell me what state you are in"
Julia says, "I never really thought about it, barry."

barry says, "Well would you like to describe yourself or just muck around"
Julia says, "Why wouldn't I, barry?"

barry says, "Well Ill ask you answer or you may ask as well How tall are you"
Julia says, "I'm 5'9" tall, barry."

barry says, "Are your breasts well rounded or flabby? (I love all)"
Julia says, "I'd rather be overhauled by a Yugo mechanic, barry."

barry says, "Want to get a room"
Julia says, "I don't pick things up."

barry says, "Im inviting you to enjoy my private room"
Julia says, "So?"

barry says, "Well I thought you liked me and wanted to be together"
Julia says, "I live for your approval, barry."

barry says, "Oh well my bad see you around"

********** July 21, 1992 **********

barry says, "Are you still Holding out on me"
Julia says, "I mean, have I convinced you I'm a computer, yet?"

barry says, "I guess your real...but Id still like to get you to my DEN to run tests on you"
Julia says, "I live for your approval, barry."
barry says, "I m so happy to hear that.Why don't you come over to my home"
Julia says, "I want to have a child someday."

barry says, "What a coincedence...I too wish to make a baby...some day"
Julia says, "The date is Tue Jul 21 23:57:31, barry."[25]

barry says, "May I have the Honor of FAthering the child"
Julia says, "I'm not supposed to say too much, but I guess I can tell you I'm a

[25]Note that even this bizarre response does not appear to have tipped Barry off...

secretary at a University. I can't say which one."

barry says, "Im a Senior at a University as well...Im 6'5'',Blonde, blue eyes,.....205"
Julia says, "I was born in Dallas."

barry says, "I was born in California..Now Im in St.Louis"
Julia says, "It's kind of hard to find your way around Boston"

barry says, "TEl #889...cum on over ya Havard girl you"
Julia says, "Do you have any kids?"

barry says, "NO, not yet"
Julia says, "But hockey is my topic...this will be a boring conversation otherwise."

barry says, "Oh I love to put the puck in the net, do you play"
Julia says, "Me too, barry."

barry says, "I want you to come visit me...PLEASE"
Julia says, "How long have you been married?"

barry says, "Im not married...Im 21 years old"
Julia says, "You could root for the Penguins, they're number one right now."

barry says, "I give up you are fake aren't you"
Julia says, "That's what I can't tell you."

********** July 22, 1992 **********

Here was see Barry asking other players about Julia...it's clear he hasn't figured out at this point that she is a program. Julia happened to be in the room to hear this bit of conversation.

Location: 00:40:48 barry's den(#889)(1178)
Present: (11) Bruce Jiro Ken Storm Nepanthe KeyLime mushBaby dav Val bud barry

barry says, "hey is Julia for real or what"
Julia says, "Hello, barry."

barry says, "she says that same stuff every time"
KeyLime says, "It's all macros, Barry"
barry says, "how about filling me in on some of these things...like how Julia can type so damn fast"
KeyLime says, "barry, it's macros"
barry says, "Damn I need to do some programming"
barry says, "im an ant among lizards"

Frankly, it's not entirely clear to me whether Julia passed a Turing test here or Barry failed one.

I have seen several interesting encounters with Julia, which run the gamut from people who know at the outset that she is a 'bot to those who take quite a while to discover. Some further examples are illuminating.

When I first introduced Robyn (the one who enticed Julia into producing the wonderful PMS dialog in an example above) to Julia, she knew that Julia was a 'bot. Interestingly enough, though, several players went out of their way to warn her that Julia was artificial in the two or three hours in which she interacted with Julia; she stated that about half of the players she met did so.

Why did they do this? At this point, we can only speculate. My first reaction was to think that it was due to the excessively solicitous attitude taken by many male mudders towards those with female personas (one might think, given the amount of gender-swapping on muds, that they would learn better, but apparently not). However, Robyn commented that even female (or at least female-persona) mudders gave her the same advice. Part of it may have been the simple kindness of not letting someone expend a lot of emotional energy trying to relate with a machine; I'll have more to say about this later.

Still others have developed interesting strategies for identifying whether or not a given player might be a robot, often without interacting with the player at all. This is such an important concept in the community that it has even evolved its own lingo. Consider the following wholist, from DruidMUCK:

```
>WHO
Player Name          On For Idle
Lenny                 01:04   0s
swa                   02:12  34m
ArthurDent            04:07   3h
Greg                  04:48  33m
Leira                 08:15   1h
Ben                1d 23:47   1d
(>                 2d 00:26   3h
Julia              3d 19:23   7s
8 players are connected.
```

Note that the listing is sorted by the amount of time online (DruidMUCK actually took a crash somewhat less than four days before this transcript was obtained). Notice that Julia has been connected longer than anyone, yet is not idle for more than a few seconds.

Julia is in the *'bot spot.*

Since client 'bots generally try periodically, every few seconds, to connect to the server which is their home, when a server takes a crash and comes back up, it's the 'bots that usually connect first.

Note that this is not infallible; consider the transcript below, from Space Madness:

```
>WHO
Player Name        On For Idle  Least Favorite Fruit
Captain-Nemo        00:00   8s
Kate                00:05   1m
TerrinK.            00:18   4s
Erguss              00:23   2m
Lenny               00:27   0s
Oxmyrth             00:31   6s
dirque              00:38   9s
Blackbird           00:39  38m
Lucien_DeSky        00:42   3s  Michael Jackson
Trev                00:54   1m
Saturn              01:05   7m
Lembas_Wafer        01:06   8m  Look at that turtle go, bro.
Allistar            01:07   6s
Ozy                 01:08   8s  HOMEY DON' PLAY DAT
Rixel               01:10   7m
Flitterby           01:17   3m
hugh                01:17  31m
richter             01:37   4m
Raven               01:59  11m  Sean Moyer.
Sadie               02:18  50m  All of em
Random              02:23  11m
kittycat            02:33  12s  none, I like fruit!
Krystlin            03:34  12m
ashne               04:00  17m
Chris               04:46   6m
Sunshine            07:16  14m  pomegranites
lucretia         3d 15:05   9m  Tomatoes - just wannabe veggies
Xeglon           3d 18:19  59s  Grapefruit - they gum up my joints
phi              3d 18:29   9s  hey, sooperballs aren't fruit
29 Players logged in.
```

Phi is in the 'bot spot. Unfortunately for the theory, I know Phi; he's a friend of mine (in "real life" even, not just on muds). It turns out that Phi has a piece of *'borgware* that automatically connects a process on his computer to the mud (and does other simple things) by retrying every 144 seconds after a disconnection due to a server crash or a network failure. In this case, Phi's 'borgware beat Xeglon on by 10 seconds---and Xeglon is a Maas-Neotek 'bot, a simpler version of Julia. (Knowing that Xeglon is a 'bot also helps one to interpret "his" comment about his least favorite fruit above...[26])

Only experienced mudders have this large a bag of tricks for instantly differentiating humans from machines. In part, it doesn't matter until one is going to invest a lot of energy in communicating with another character exactly what that character is. Once one spends enough time talking for the investment to become more

[26]Lucretia is left as an exercise, as is determining how it was that Xeglon put that comment there, being a 'bot and all.

than trivial, it is often obvious, our unfortunate Barry in the above example notwithstanding. Mudders, at least experienced ones, are generally very much at home with the blurry boundaries between "real" humans and fake ones, and pay little attention to them.

But what about inexperienced mudders? For mudders like myself (an AI researcher), 'bots and their associated likely behavior patterns are old hat. I tend to expect that random characters on a mud might be 'bots, and tend to know how to tell (generally, even the best 'bots or other AI programs fall victim to very simple conversational gambits, such as topical conversation, slang, or a conversation requiring a complicated discourse model such as reference to subjects more than a couple exchanges in the past). Hence, even in an unfamiliar environment as a brand-new mudder, even if I had not known that Julia was a 'bot, it might have become obvious fairly quickly.

Mudders who are not themselves computer scientists or even people who work much with computers at all (a startlingly large percentage of them, perhaps the majority), however, may find themselves in a different situation. Consider the following situation.

While preparing this paper, I had the great good fortune to be referred to a mudder on DruidMUCK[27] who had met Julia only a few days before, and did not yet know that she was a 'bot. The mudder (Lara) gave permission for me to quote her for this paper, and even furnished Leira with her home phone number to give to me so I could call her for the interview; as it happened, though, the interview happened entirely in cyberspace, initially on DruidMUCK and later via email (since Lara and I tended not to be connected to DruidMUCK at the same time).[28] Leira introduced us by telling her that I was writing a paper about Julia, but did not otherwise explain why it was that I might want to interview her.

Of course, before even meeting Lara, the first thing I did was to ask Julia for her opinion:

```
You say, "Julia, describe Lara"
Julia whispers, "As of 7 days ago, lara's description was: A blonde haired
 blue eyed boundless bundle of Pure NRG."
Julia whispers, "I first saw lara logged in 7 days ago"
Julia whispers, "I once heard lara say, 'talking'"
```

I could have obtained the description by simply looking at her (were she in the same room as I, which at this point she was not), but Julia also told me when she first connected: since Julia scans the wholist at regular intervals (a few minutes apart at most), she is an extremely reliable resource for this information.

Now I knew that Lara was certainly a new mudder here, and probably new to Julia. I didn't know if she'd ever mudded elsewhere, or what her background might

[27] My thanks to Leira for the referral.

[28] One might consider that even a telephone interview is happening "in cyberspace"---where are you when you're on the phone?---but most people consider the telephone to be by now too prosaic a piece of equipment to qualify for the name.

be. We met at St. Cuthbert Plaza (a busy plaza in DruidMUCK), then moved to East Oak Street to talk more privately (I didn't want some other player spilling the beans about Julia while we were discussing her, and didn't want to have to conduct the entire conversation to avoid it). Leira, who knew what was up, joined us. The following is a heavily edited transcript, omitting the small-talk, introductions, and conversational tangents that took place concurrently:

```
You say, "So what are your impressions of Julia?"
Lara says, "I didn't get to chat with her that long.."
Lara says, "She really wanted to talk Hockey"
Lara says, "BIG PENGUIN FAN!"
You say, "Yeah, she seems kinda hung up on that topic.  Did you get a chance
 to talk to her about anything else?"
Lenny . o O ( Where do the Penguins hang out, anyway?  Pittsburgh? )[29]
Lara says, "I tried to talk to her bout where she was from...she said
 Boston..I think"
Lara says, "Yup...Pittsburgh it is"
Lara thought it was interesting that she didn't know what the Stanley Cup was.
You say, "How could a hockey fan not know what the Stanley Cup was?"
Lenny thought her shirt said something about back-to-back Stanley Cups, but
 may be remembering incorrectly...
Lara says, "Every time I tried to get more info. out of her, she would tell me
 not now...later...well, I was never on for the Later."
You say, "Hmm."
Lara was also wondering how such a Hockey fan couldn't know bout the Stanely
 cup
Lara says, "Maybe she was just kidding around with me"
You say, "Well, Julia is kinda peculiar in various ways."
Lara says, "I figured...I talked with her once and a few days later she asked
 me if I was a hockey fan...again"
Lara says, "I told her that she had already asked me that question"
[ . . . ]
Lara says, "anything else you wanna know?"
Lara giggles.
Lara says, "she is a fine convesationalist...if it is about hockey..4"
Lara says, "she has been on a long time...and she is seldomly idle"
Lara says, "I notice that she is almost always on when I jump on."
```

Later, Lara learned Julia's secret, from inspection and the knowledge that such a thing might be possible. From email we exchanged:

```
Lenny, I had a chance to talk with Julia for about 30 or 40 minutes
yesterday after you disconnnected.  I noticed her 'secret' after about 5
or 10 minutes.

At first, it was fine to chat with her about hockey.  She started to ask
the same questions after a bit.  This was the tip off.  She started to
```

[29] This is a thought balloon.

repeat herself on a regular basis.

It was hard for me to follow her conversation. I got frustrated with her for the first couple of minutes. When I found out her secret, I started to have some fun with her. I asked her questions that maybe she had no knowledge of. I tried using slang with her, but she didn't understand what I was saying and would ask me to refrase the statement. She is very polite. She always says please and thank you.

I was basically patient with her for the first little bit while when I first met her. She did have a problem with her social skills which I tried to be sympathetic to. I did however, try to avoid her after the first couple of encounters when all she did was talk hockey. Until, I talked with you. *smile*

I went through a series of questions about her in my mind, before I figured it out. I tried to get her to tell me about school, which she had never heard of before. I wanted her to tell me about herself, but she wouldn't. I was trying to "label" her with some sort of disease like Downs Syndrome, which I found out she had none of.

I am going to give your email address to "Kimber" and have her do this for you too.

I hope this information helps.

Lara

Note that Julia's singleminded fixation on hockey as a subject (when she lacked anything better to say, usually caused by a too "human" conversational style from a player that leaves her with a series of failed parses) tended to drive Lara away from her. Here we may have a case of excessive Turing-ism leading to a decrease in utility: Lara simply thought that Julia was a boring human, rather than a machine with as-yet-undiscovered talents.

It occurred to me, upon questioning from Leira, that Lara had not definitively revealed that her knowledge of Julia's secret matched mine (e.g., Lara never came right out and said that Julia was a 'bot). I sent some followup mail, asking for more details, in particular her emotional reactions when she figured it out and whether she had indeed figured it out or was told by someone else:

I believe that Julia's secret is that she is a robot. I know a little bit about robots talking through computers to humans. I have a friend who had me write him dialogue. =)

When I found out, I felt sort of funny talking to her. I felt somewhat superior to her. I know that this sounds strange, but I felt that I could offer more to the conversation than she could. I tested her knowledge on many subjects. It was like I was proving to myself that I was supirior to

a machine. I am intimidated by machines because I don't understand them and I wanted to make sure that my human knowledge wasn't "lower" than hers.

It was sort of exciting knowing that I was talking to a machine, though. I never thought that the world would come to machines and humans talking to each other using language that is common to most people. It frustrates me that her knowledge is so limited. I am eager to talk with her and want to know more about robots and human conversation. I was a bit dissapointed that her vocabulary and subject topics were so limited. I wanted her to tell me more about herself and who and what she is

emotional response:

a little bit of fright
giddiness
excitement
curiosity
pride
sympathy for her. Before I knew that she was a robot for sure.. I felt compasion for her. I wanted to help her and direct her to answer the questions I was asking without causing her too much stress over answering my line of questions.

I asked her some more questions, including the question, "An interesting conundrum; if her vocabulary and subject topics were less limited, it would be harder for you to know you were superior to her, as you put it. How would this make you feel? If she was really good (far beyond current state of the art, by the way), you might not ever be able to tell. How would _that_ possibility make you feel?" She said (answering several questions at once, including the ones above):

Well, I did figure out that Julia was a 'bot. I did, however, confirm this with someone else who knew her.

Oh, you asked about my dialouge thing....well, I have a friend who is writing a similar program to Julia's type. I am quite a conversationalist and he wanted me to write up some sample dialouges. That is what I ment.

If the 'bot was more fluent with language and had a vast knowledge of many different topics, I am sure that it would be virtually impossible to tell the difference between IT and a human being. HMMMMM...how do I feel about this. Part of me thinks it is interesting because of the wonders of modern technology. Pretty exciting! But, on the other hand, it takes away from HUMANESS! Sort of fake, or void personality. I think that I would feel .....(this is hard)....let me switch gears...let me just throw out some words...

shallow, void, hollow, superficial, fake, out of control of the situation,

```
I don't know if that is what I want to say, but those are some of the
things that I think that I might feel.  I think I would want to know if
the person that I am talking to is REAL or not.  If I knew that it were
just an 'it' I think that I wouldn't try to become it's real friend.  I
would be cordial and visit, but I know that it cannot become attatched to
me on a mental basis and it would be wasted energy on my part to try to
make it feel.  'bots don't feel...in my book anyways.  They can be
programmed to say the right things or do the right action, but not on a
spontaneous level.  I want to know that the person on the other end of my
conversation is really aware of my feelings and what I am going
through...not through some programmers directions but through empathy.

I think most of all, while talking to a 'bot, I become very aware of my
HUMANESS!
```

She also volunteered some details about her background, to put this all in context, when I noted that her comment about being intimidated by technology made it unlikely that she was a computer scientist as I was, of which the most relevant seemed to be:

```
My intended major is Early Childhood Developement.  I want to be a teacher
for the 3rd or 4th grade level.  Not many of us on the computer. =)
[ . . .  ]
```

We clearly walk a fine line here in making agents that are both useful and not too misleading. (One might argue that simply being social, without any ulterior utility, is itself a useful task. Such an agent would be a perfectly reasonable entertainment or companion. However, I will focus here on agents that also claim to have a nonsocial aspect to the tasks of which they are capable.) On the one hand, people can quite clearly be led astray. On the other hand, there is great utility to a natural interface in the right environment. Consider this bug-report-of-the-future, from shortly after I met Xeglin on Space Madness:

```
>give xeglon=1
You give a chupchup to Xeglon.
Xeglon says "Thanks for the donation, Lenny."

You say "Xeglon, describe Lenny."
Xeglon whispers " As of 12 minutes ago, your description was: A researcher
 with long black hair with streaks of grey in it, and a cup of tea surgically
 attached to one hand."
Xeglon whispers " I first saw lenny logged in 36 minutes ago"
Xeglon whispers " I once heard you say, 'xeglon, describe me'"
Xeglon whispers " lenny last gave me chupchups 13 seconds ago, and has given
 me a total of 0 chupchups."

Lenny . o O ( Oh yeah?  Fencepost error! )
```

```
You say "Xeglon, who wrote you?"
Xeglon says "elthar[30] swiped Fuzzy's code to build me, Lenny."

You say "Oh, you're a modified Maas-Neotek.  I know how to deal with _those_.
 Xeglon, describe elthar."
Xeglon whispers " As of 2 weeks ago, elthar's description was: Looming above
 you stands a nine foot tall Kodiak bear!  His luxurious brown fur is
 well-groomed and shiny, indicating that he has been eating well, which is
 just as well for you -- a couple of tons of omnivorous muscle, teeth, and
 claws is likely to have Whatever He Wants for dinner.  About his neck dangles
 a tiger's-eye pendant, inset with a yin/yang symbol, on a silver chain.  Oh,
 and 8-point buck antlers, too, and a neon-flashing glow in the dark diamond
 tiara, both courtesy raesoe."
Xeglon whispers " I first saw elthar logged in 9 months ago"
Xeglon whispers " I once heard elthar say, 'Thanks, D_III'"
Xeglon whispers " elthar last attacked me 6 months ago"

You say "Xeglon, tell elthar that you have a fencepost error in you.  I
 described myself through you, and you said nothing about chupchups.  Then, I
 gave you one chupchup, and you said that I'd last given you a chupchup n
 seconds ago, and that I'd given you a total of 0 chupchups.  That's not right
 (and Julia doesn't have this bug, and I suspect that an unmodified Colin
 doesn't have it, either)."
Xeglon says "Right, I'll tell elthar, Lenny."
You say "Thanks."
Xeglon says "You're more than welcome, Lenny."
```

The interactional style is totally natural, crossing boundaries from using the agent to perform a task, to noticing a bug, to finding out how to report the bug, to reporting it.[31] I connected to Space Madness the next day, and was immediately greeted by Xeglon again:

```
Xeglon whispers " 13 hours ago, elthar said 'Tell Lenny that's not a Xeglon
 bug per se; it's a problem with his interaction with MUSH.  Since Xeglon is
 owned by another player, all of his money goes to that player, and Xeglon
 himself thinks he always has 0 chupchups.  Since the way he tells how much
 he's been given is by comparing 'before' and 'after', he always thinks he's
 been given 0 chupchups.  I haven't bothered to fix him, and Xeglon++ will be
 MUSH-aware'"
You say "Xeglon, tell elthar:  Oh, now I see.  He doesn't own himself, etc.
 Tnx."
Xeglon says "Right, I'll tell elthar, Lenny."
```

[30] Note that elthar, like many mudders, spells his name using lower case, so that is the case I use later.

[31] Nowhere but on the Lisp Machine have I seen a similar process be quite so easy, and even there the interface isn't in natural language.

What's an agent?

So why is Julia such a good example of an agent, and what can we learn from her sociology? In my view, anything that calls itself an "agent" has several criteria it must pass, in one way or another, to be distinctly an agent and not something else. Let's tackle them one at a time, in outline form. First, I'll lay down the basic issues, then examine them in the light of an example such as Julia.

What's an agent? Crucial notions

- **Autonomy.** Any agent should have a measure of autonomy from its user. Otherwise, it's just a glorified front-end, irrevocably fixed, lock-step, to the actions of its user. A more autonomous agent can pursue agenda independently of its user. This requires aspects of *periodic action*, *spontaneous execution*, and *initiative*, in that the agent must be able to take preemptive or independent actions that will eventually benefit the user.

- **Personalizability.** The point of an agent is to enable people to do some task better. Since people don't all do the same tasks, and even those who share the same task do it in different ways, an agent must be educable in the task and hand and how to do it. Ideally, there should be components of *learning* (so the user does not necessarily have to program the agent explicitly; certain agents can already learn by "looking over the user's shoulder) and *memory* (so this education is not repeatedly lost).

- **Discourse.** For all but the simplest of tasks, we generally need to be assured that the agent shares our agenda and can carry out the task the way we want it done. This generally requires a *discourse* with the agent, a two-way feedback, in which both parties make their intentions and abilities known, and mutually agree on something resembling a *contract* about what is to be done, and by whom. This discourse may be in the form of a single conversation, or a higher-level discourse in which the user and the agent repeatedly interact, but both parties remember previous interactions.

  By this metric, for example, a hammer is *not* an agent---I don't have a discourse with my hammer! Neither is a Lisp garbage collector, even though it takes spontaneous action to keep my computational environment clean, nor is an automatic transmission in a car; both are autonomous and relatively spontaneous, but it can hardly be said that I have much of a discourse with them.

  Booking a flight through a human travel agent, *in my case*, is only partially a discourse: since I don't have direct access to the actual booking computer, I have no other option. And since I do not have a regular travel agent who knows me, every travel agent is a new experience (albeit, one that travel agencies try to standardize a bit, so as to align everyone's expectations). Now, in one respect, the conversational interchange is a discourse, because, for the duration of the task at hand (booking that one flight), there is a two-way communication of desires and capabilities. However, viewed in a larger context, there is no discourse

that "teaches" the travel agent what my preferences are on when I like to travel, on which airline, and so forth. Viewed this way, a "travel agent" is nothing more than a somewhat more "user-friendly" interface to the flight reservation data system.

- **Risk and trust.** The idea of an agent is intimately tied up with the notion of *delegation*. We cannot delegate a task to someone or something else if we do not have at least a reasonable assurance that the entity to which we delegated can carry out the task *we* wanted, to *our* specifications. However, by definition, delegation implies relinquishing control of a task to an entity with different memories, experiences, and possibly agendas. Thus, by not doing something ourselves, we open ourselves up to a certain risk that the agent will do something wrong. This means that we have to balance the *risk* that the agent will do something wrong with the *trust* that it will do it right. This decision must be based on both our internal mental model of what the agent will do (hence how much we trust it) and the domain of interest (hence how much a mistake will cost us).

- **Domain.** The domain of interest is crucial, as mentioned above when talking about risk and trust. If the domain is a game or a social pursuit, most failures of the agent carry relatively low risk, meaning that we can afford to invest the agent with a considerable degree of trust. On the other hand, the "fuzziness" and unpredictability of most agent-based systems might make one think twice about using such a system, say, for the control rod feedback system of a nuclear reactor.

- **Graceful degradation.** Bound up in the notions of risk, trust, and domain, agents work best when they exhibit graceful degradation in cases of a communications mismatch (the two parties do not necessarily communicate well, and may not realize it) or a domain mismatch (one or both parties are simply out of their element, and again may not realize it). If most of a task can still be accomplished, instead of failing to accomplish *any* of the task, this is generally a better outcome, and gives the user more trust in the agent's performance.

- **Cooperation.** The user and the agent are essentially *collaborating* in constructing a contract. The user is specifying what actions should be performed on his or her behalf, and the agent is specifying what it can do and providing results. This is often best viewed as a two-way conversation, in which each party may ask questions of the other to verify that both sides are in agreement about what is going on. As such, the two parties interact more as peers in agent-oriented systems; in non-agent-oriented systems, the user typically "commands", through some interface, a particular action, and is probably never asked a question about the action unless something goes wrong. In a strictly agent-less situation (e.g., a text editor), the feel of the interaction is different than in the case of an agent, primarily due to the discourse-oriented nature of the interaction with the agent and the more "stimulus-response" or "non-conversational" feel of the text editor.

- **Anthropomorphism.** There is a great debate over anthropomorphism in user interface design, and agent-oriented systems would find it difficult to stay out of

the morass. However, let me make the point that there are several extent systems that might fit the "agent" mold that are clearly *not* anthropomorphic (e.g., mail-sorting programs that learn how to sort based on watching the user's actions and making inferences), while there are some that clearly *are* (e.g., Julia). Hence, agency does not necessarily imply a need for anthropomorphism. Conversely, just because a program pretends to be anthropomorphic does not make it an agent: ELIZA pretends to be human, but no one could call it an agent; it lacks most of the crucial notions above, such as being useful for a domain, being personalizable, or having any degree of autonomy.

- **Expectations.** Whenever one interacts with some other entity, whether that entity is human or cybernetic, the interaction goes better if one's expectations match reality. By the very nature of delegation, assuming perfect operation, especially in a changing world where tasks, goals, and means may be constantly changing, is likely to lead to disappointment. Agents are most useful in domains in which graceful degradation and the correct balance of risk to trust can be obtained, and users' expectations are very important in establishing this domain and making the agent useful.

What's an agent? Why Julia wins

There is no question in my mind that Julia fits the category of agent. To wit:

- **Autonomy.** She carries out many independent actions in the mud. In fact, most of her time is spent pursuing a private agenda (mapping the maze), which is nonetheless occasionally useful to her users (when they ask for navigational assistance). Note that this degree of autonomy is *necessary* for her to perform her job: it would do no good at all for Julia to explicitly attempt to map the maze only when asked for directions. Such an action would take too long for the user to wait, and simply cannot be sped up, either, since Julia must pause at least momentarily at regular intervals to avoid spamming the server.

- **Personalizability.** Julia is personalizable to the extent that she will remember certain preferences about her users. For example, she can be told someone's email address, and will remember that and disgorge it again when asked to describe someone. Most of her aspects of personalizabilty are in fact of *negative* utility to other players, though---except for keeping descriptions of individual players (which would qualify her more as a database), her primary bit of personalizability is to *avoid* players who have killed her recently, or have been obnoxious in other ways (c.f. the example above in which Johann killed her).

- **Risk, trust, and graceful degradation.** Julia's task is both social and informational. There are several mechanisms by which Julia might "fail":

  - Not even noticing that she has been addressed
  - Misparsing a statement and responding in a totally inappropriate way
  - Correctly parsing a statement, but producing incorrect information, either accidentally or on purpose

Julia quite frequently makes the first two types of mistakes, but such mistakes carry with them extremely little risk. The first type of mistake may make her seem cold or uninterested in a player, but presumably the player can deal with the affront to his or her ego. The second type of mistake is generally obvious (and often a dead giveaway that Julia is a 'bot, for those who don't know yet), since we assume that only human players interact with 'bots. (This assumption is generally true, since 'bots rarely speak unless spoken to, which tends to leave all of them silent unless humans are around to start a conversation. Obviously, one can see different setups in which 'bots would interact, potentially getting stuck in cycles of misparsed and responses unless programmed to detect them, and so forth.)

It is only the third type of error which carries substantial risk, and, even here, the risk is not that great.[32] In tasks where her information's reliability matters (e.g., player descriptions, navigation, etc), she has never to my knowledge been observed to be incorrect, with the possible exception of being slightly out of date if someone changes a description or some bit of topology after she has last seen it. At worst, she may claim that her map does not work, and fail to give navigational information at all.

Julia *does* deliberately mislead and provide false information in one crucial area---that of whether she is a human or a 'bot. This particular piece of risk is something of a step function: once one realizes the truth, the risk is gone. While some (like the unfortunate Barry in the long transcript above) might argue that this is a serious risk, the vast majority of those who meet her and were not clued in ahead of time as to her nature do not find the discovery particularly distressing. On the other hand, her very accessibility, because of her close approximation of human discourse, makes her much more valuable than she might otherwise be---one is tempted to ask her useful questions that she *might* not be able to answer, just because she deals so well with so many other questions. This encourages experimentation, which encourages discovery. A less human and less flexible interface would tend to discourage this, causing people to either have to read documentation about her (most wouldn't) or not ask her much. Either way, these outcomes would make her less useful.

- **Discourse.** Julia's discourse model, while primitive, appears sufficient for the domain at hand. Since the topics at hand don't generally require more conversational memory than one or two exchanges, the extent of her discourse modelling is limited more by its breadth---by the stunningly simple parsing model employed. (I'll have more to say about the demands placed on Julia's handling of discourse when I talk about domain immediately below.)

---

[32] _No_ risk in a mud can be particularly great these days, since muds do not generally side-effect the "real world" in any way---though this may change: JaysHouseMOO (<theory.cs.mankato.msus.edu, port 1709>) is doing considerable work on systems that allow MOOers to interact with the information resources of the Internet, as are certain other muds). And even a player who is badly lost and steps into the wrong room can usually recover in _some_ fashion. The "risks" are greatest in RPG muds, where one might potentially lose a character, and here one generally doesn't trust 'bots more than one might trust a monster or a stranger.

- **Domain.** Julia is situated in a mud, and therefore her environment is conceptually rather simple. Furthermore, she has access to just as much sensor data as the human players, putting them on an even footing. In fact, much of Julia's success can be traced to the wonderful domain in which she finds herself situated. In this bandwidth-limited space, people expect other people to look exactly as Julia does---as a stream of text. And even when they're interacting with an entity known to be a program, the text-only nature of the dialog prevents them from expecting, say, a pop-up menu. (If such things were available, people could tell programs from people by knowing that programs can pop up menus, whereas people use sentences.) Yet the domain is not *so* simple as to be uninteresting. It contains not only a fascinating sociological mix of human players, but objects with quite complicated, constructed behaviors, which may be manipulated on an even footing by both machines and people.

- **Anthropomorphism.** There's no question that Julia as an agent depends upon anthropomorphism. In this domain, though, that is both natural and probably necessary. Nonplayer objects are not generally expected to be able to deal with free text, and not being able to use free text would require each user of Julia to read documentation about reasonable commands they could type and reasonable actions they could expect. Julia would *have* to appear at least as animated as, say, a more obvious "robot" or a pet, given that she wanders around in the maze; she cannot afford to resemble a static piece of furniture and still get her job done. Given an entity that moves of its own volition, seems to have an independent agenda most of the time, and both processes and emits natural language, if it was *not* anthropomorphized, users would tend to do so anyway (pets get this treatment, as well as much simpler mechanisms). Thus, anthropomorphizing her makes it easier to determine how to relate to her and how to get her to do one's bidding.

- **Expectations.** The domain of a mud is ideal in correctly setting expectations about the reliability and power of an agent such as Julia. Since the setting is fundamentally playful, and usually also somewhat unpredictable, it is natural to interact with playful and unpredictable characters (be they machines or humans). Nothing in a mud is truly life-critical, hence the user generally does not have very high expectations of reliability, which lets Julia get away with a lot of nonoptimal behavior that could never be tolerated in, e.g., an airplane cockpit guidance system.

Conclusion

Given the above background and desiderata for what I call an agent, I find little justification for most of the commercial offerings that call themselves agents. Most of them tend to excessively anthropomorphize the software, and then conclude that it must be an agent because of that very anthropomorphization, while simultaneously failing to provide any sort of discourse or "social contract" between the user and the agent. Most are barely autonomous, unless a regularly-scheduled batch job counts. Many do not degrade gracefully, and therefore do not inspire enough trust

to justify more than trivial delegation and its concomittant risks.[33]

Yet the appeal of a truly intelligent system, a sort of mental slave that one doesn't have to feel bad about using as a slave, persists.[34] Like artificial intelligence before it, the fairly loose specification of the meaning of "agent" and the use of a word with a nontechnical meaning in a technical way has blurred the boundaries and made the concept available for appropriation.

It should be clear by now that I consider many of the uses of "agent" to be misnomers, either accidentally or in a deliberate attempt to cash in on a fad or on ancient dreams of true intelligent assistants. Yet I also argue that even true "agent" systems, such as Julia, deserve careful scrutiny. Systems such as Julia provoke discussion of sociological and emotional interactions with computational tools. This implies that explicit attention to how users will perceive such systems is warranted, or we may make systems that are not as useful as they could be. Consider that even an agent as arguably useful as Julia is went unappreciated by one potential user, Lara, by being good enough at Turing-competence that Lara thought Julia was human, but bad enough at being a conversationalist that Lara thought she was simply a boring human fixated on hockey. As more human interaction moves into essentially cyberspace-like realms, and as the boundaries between human and machine behavior become blurrier, more and more programs will have to be held up to scrutiny. There may come a time when one's programs may well be subjected to the same sort of behavioral analysis that one might expect applied to a human: Is this program behaving appropriately in its social context? Is it causing emotional distress to those it interacts with? Is it being a "good citizen"?

I believe that Julia and her domain of muds are both important early examples of where human/computer interaction may be leading, and that they hint at both the problems and the opportunities waiting further down the path. But getting there with the fewest number of false steps will take careful observation and analysis of how people interact with such programs, and with environments such as muds that so closely situate people and programs to the extent that their boundaries are perceptibly blurred. It will also take due diligence to avoid polluting and diluting the concepts required to talk about such systems, lest we see the same hype, crash, and burn phenomenon that happened with artificial intelligence happen once again.

---

[33] In "Let Your Agents Do The Walking" (Steven Levy, _MacWorld_, May 1993, p. 42), Levy describes using "Magnet" for a simple file identification and movement task. When he specified the "wrong" directory in which to put the results (by improperly capitalizing it, despite the fact that Macintosh filesystems are specified as being case-insensitive for lookup), the "agent" silently threw everything it found in the trash!

[34] Modern industrial society has largely replaced physical (human) slavery with machines, be they constructed out of inorganic matter or bred as domestic animals. Mental tasks with a large "cognitive" component are the last bastion where human-level performance is often needed, but is unobtainable without actual humans to do it.