Early Programming

by Sara Roberts

MFA Thesis - Spring 1989
California College of Arts and Crafts

This is dedicated to my grandmother, Maude Killam Becker,
who strongly believed in good manners, proper grammar, and modern art.

Acknowledgments

I would like to express my fond gratitude to Jim Crutchfield for his ideas and insight, for listening, considering, discussing, and answering my endless questions. I thank him for both his remarkable patience and impatience. His involvement and support have immeasurably improved and enlivened every aspect of this project, and it wouldn't have worked without him.

I also thank the members of my committee, particularly Donald Day who has been my mentor and incomparable friend throughout an extended graduate career. Jody Gillerman and David Heintz have been valued friends and advisors, as well. In addition I would like to thank Lynn Kirby, Lynn Hershman, and Steven Goldstine for their suggestions and encouragement.

For generously lending and helping to obtain hardware for this project I thank Wayne Wagner of Lucasfilm, Ltd. and Norman Packard at the Center for Complex Systems Research, University of Illinois, Champaign-Urbana.

Finally I would like to thank my parents for their hand in my own early programming and their support and encouragement ever since.

Things

> "It is necessary to be quite clear about the universal truth that whatever 'things' might be in their pleromatic and thingish world, they can only enter the world of communication and meaning by their names, their qualities and their attributes."
>
> - Gregory Bateson, **Mind and Nature**

> " The notion that the reasons for things derive from a human will, or from a supernatural analog of human will, expires rather late in the life of each child. But even you and I will occasionally get angry and kick our automobiles and yell at them, as if we had regressed to the primitive way of thinking of objects as if they were alive."
>
> - George Miller, **States of Mind**

Life as a human being seems to be in great part a matter of interacting with objects, things that do not notice when you touch them, things that we generally agree don't have a life of their own. But in fact it seems useful, inevitable for us to animate them, attribute qualities of life to them. We don't do it consistently or particularly consciously, just when, for instance, the thing crosses us, when it (willfully) stands in the way of our progress. Or when it has helped us, surprised us, pulled us through - then a thing can be a friend, ally, and comrade, however fleetingly.

A thing with some degree of unreliability seems particularly magnetic. Machines, for some reason seem to draw our animation much more readily than simple objects like cups, rugs, or eyeglasses. A can opener, for instance, can harbor much more malice than a can. And the more complex the thing the better. Cars are good candidates. Our relationship with them is reinforced with such dependence, their workings and not workings have such consequence.

I'm curious about this. Does being able to identify with an object help us or fool us? Why is it we are so generous with attributes of character? Maybe this generosity helps us extend ourselves into the rarefied air of abstraction, like some kind of bootstrapping, a leg up into symbolic thought. Or is it primitivism, using emotions where emotions are not.

Further Aspirations, a video piece I made in 1981 extends that animating tendency to its logical extreme - it portrays people romanticizing the lifestyles of objects, wanting to be like them, wishing they could turn into these things they admire. It begins with a young boy proudly proclaiming his intention to grow up to be a jet, not just any jet, but the Concorde. Next a high school girl reluctantly tells her counselor that she'd like to be a Coleman cooler, or lantern maybe, because they're there when you need them in adverse circumstances, like heroes. (The counselor reminds her of the many months she'd be spending in the basement with the Christmas tree ornaments.) A young man then describes joining the incredibly exciting worldwide intelligentsia we know as telephones ("... beneath that sleek plastic exterior lies an international force! "), a dishwasher speaks of his fondness for words and his desire to become the Oxford Unabridged English Dictionary, and a fashion model wishes she could just be a picture of herself, (maybe a Scavullo or an Avedon).

We invent meaning to relate ideas, events, and images. We make stories or narratives out of the jumble of events. We see patterns we call style or character, we find personality. More projection.

As an example, I remember that when I was first learning arithmetic there were some numbers I liked and some I didn't and my feelings about them would shift, but each number had a definite character. There was a smug fullness to two, three was kind of charmingly geeky, five had a paternal orderliness to it that was magnified in ten but made gross in one hundred. Seven was mysterious and adult, one very honest and pure, but homely. Zero a shifty cheater and nine a social climber, always aspiring to the decade above it. I only dimly remember this early acquaintance with numbers, but these first impressions must be still in some way operative in my mind. I wonder if the attributes I have given numbers get in the way of understanding the pure quality of quantity, (I'm not very good with arithmetic), if the mumble of interaction between these 'characters' distracts me. Or maybe character is essential, and to project these traits into numbers is necessary in order to work with them, understand them, have an impression of them at all. How does one correctly understand a number? Do feelings have anything to do with it?

Strangely enough I don't have any similar sense or remembrance of attaching character to the alphabet, letters seem sub-animate to me. Of course as soon as they combine to

become words, take on meaning, they have character as a by-product. And, besides the public meaning of a word, there is a sort of private history with some words. (The first time I heard that, etc.) Then there is type style, the shape of the ink on the page, which gives each letter a kind of objectness. And having that we can sense a certain character, a phantom voice, familiar to us, providing context. This is the newspaper voice, the grade school textbook voice, the classical literature voice, the amateur voice of the typewriter rapidly being replaced with the amateur voice of the dot-matrix printer, soon to be superseded by the ubiquity of the laser printer.

Helvetica, a video piece from 1980, is about the typestyle of the same name....that international forward-looking but humanistic voice, a little passé now, but still very popular in public. Again, this tape takes the idea to an instinctually logical, but questionable conclusion - where there's a voice, there must be an intelligence. Text, handwritten and typeset, runs across the screen introducing the character, Helvetica, trying to distinguish its particular identity. This gives way to a dream, finally, pursuing the phantom voice through public space.

Where there is a voice, where words are spoken, written, even recited, it's natural to believe that there is an intelligence behind them. We think that at some point, somewhere, a person wrote that, said that, thought that. A message is coming to us from ... someone?

Now, as I sit comingled with the interface design of my personal computer, an extended voice guides my every move.

Artificial Intelligence

> "It is true that artificial intelligence may go beyond printouts into artificially voiced speech. It may move beyond printing to the more subtle embodiment of meaning that occurs in sounds. If it succeeds in doing so, it's 'speech' will have been a transformation of its writing and will bear the imprint of writing. Artificial intelligence will have moved in a direction that is the reverse of that followed by natural thinking, which went from voiced speech to the written word."
>
> - Robert Sokolowski, **Natural and Artificial Intelligence**

This is where the computer got interesting for me. A thing that traffics in words. A perfect object to attract our animation, just the sort of thing we like to make 'real'.

Years ago, maybe 1979, I picked up a book called "Artificial Intelligence and Natural Man" by Margaret Boden, a British philosopher.[1] It was then and still is a most level-headed survey and critique of artificial intelligence. The book details a period when many people in artificial intelligence looked up from a wallow in grand illusions and a largesse of funds and realized how very slight their purchase on the field of intelligence actually was. Lots of them bailed out for the study of linguistics, perception and philosophy. This all happened just for a brief moment before venture capital made wallowing and hollering irresistible again.

The chess-playing, equation-solving, oil-prospecting type programs that started the field of Expert Systems are not what got me excited, though they are certainly the most likely to have some real life application. The programs that interested me were the ones that <u>really</u> bit off more than they could chew; programs attempting to simulate particular world views, programs that "thought" in characteristic ways and functioned under the governance of "belief systems". These, with their arrogant assumptions about How We Are caught my interest. They included programs simulating a right-wing ideologue, a paranoid psychotic, a Rogerian therapist, and a neurotic who believes she is descended from royalty. Then there were some with programmed beliefs that had more to do with parts of speech or the possible combinations of red and blue blocks than unhealthy passions. And these, too, in the conceits of their constructed worlds held a certain charm.

There were two in particular, both written by psychologist K. M. Colby, which attracted me. I have to make clear at the outset here that I consider these programs extremely dubious as science. But, I thought they had some interesting possibilities as forms of representation, or fiction. In short, had the nature of their pretensions been different they might have made good art.

---

[1]Strangely enough I found it at Shambala Booksellers, a long established Berkeley bookstore of an anti-rationalist bent.

4

The first was a program that, surprisingly, does not have a name (virtually all these programs are graced with cute acronyms), it's simply referred to as The Neurotic Program. This program attempts to simulate the thinking processes of a young woman engaged in dialog with a therapist according to a classical Freudian model. This neurotic had a "Belief Matrix" that included such beliefs as: People ought not to abandon people, I must love father, Father abandoned me, I hate atheists, etc. Each belief had a relative charge within a range +5 to -5, depending on how close to the "core conflict" it is. The core conflict is, of course, hating her father.

The Neurotic program starts up by choosing a belief, something like, "I must love Father." Then it will find all her other related beliefs, anything else having to do with Father, or with loving, thus forming a "Complex", which is then tested against the introduced belief for conflict. Conflict (determined by the aforementioned relative charges) will, in high scoring cases trigger Defensive Thinking, which brings in a set of Transforms used to alter the belief and make it more acceptable. She can't bring herself to say "I hate Father", but she can say, instead, "I hate my boss." If the conflict is not too serious, say, "Father is not always thoughtful", it will cause a rise in the Anxiety Level of the program, making it more susceptible to conflict the next time around. The Anxiety Level will just continue to mount until the anxiety-causing belief can be expressed, which can only happen if one of the Transforms can make the difficult thought more acceptable.

So the world of this program is a sort of tormented cycle in which any new "belief" has to be squared with all her other beliefs, with resulting small changes in the entire belief system. Reading about it I wondered if it was possible for the program to be "cured" of her neurosis. The book did not say. I imagined it being a task similar to solving a Rubic's cube, gradually wriggling all the conflicted beliefs out of the poor thing.

The emotional life portrayed here is transactional in a way far surpassing the wildest simplifications of any pop psychology. Typing in a sentence such as 'I must love Father', assigning a value to it falling somewhere in a scale of values, and calling it a "Belief", is what Drew McDermott, a computer scientist at MIT[2] bemoans as "wishful mnemonics".

_____

[2]In his essay, "Artificial Intelligence Meets Natural Stupidity" collected in **Mind Design**, John Haugland, editor.

In other words calling this a "belief" is projecting a complex state of affairs onto a simple number. And to assume that the various complex states that you have attached to these values do actually interact according to the rules of numbers is a very large assumption, indeed, in a field where assumptions are not welcomed.

What I like about The Neurotic Program is the thing that makes it least viable as science, that is (though Colby surely thought otherwise) the choices made in the representation of this personality and the values assigned to them strike me as being very bound up in cultural conditioning, a reflection of the programmer as much as his subject. A kind of portraiture. According to the book the program was designed from the case history of a particular patient of Colby's, and while he is certainly after a general model of neurosis it ends up being more expressive than objective.

Another of these programs, richer in character, is Colby's simulation of a paranoid psychotic, PARRY. PARRY's world revolves around his pastime of betting on the horses, which gets him in trouble with dishonest bookies, and maybe even the Mafia. This is where his paranoia enters the picture. It is possible to converse with PARRY via teletype, and though a conversation may start innocently enough, PARRY is scanning every sentence for cues to one of his "flare topics". Should he find anything his list of synonyms tells him is related to horses, the law, Italians, or doctors, he will go into a delusional mode focused on the Mafia. His mood can range from mildly surly to intensely suspicious and hostile depending on the presence of key words in the conversation.

Colby intended using these programs to train therapists. Each of these programs was written with the intent of giving a dynamic picture of a general condition, but seems much more effective to me as a particular characterization, enhanced by both the possibilities and limitations of its medium. In short it makes better fiction.

Another program which I found fascinating, not for it's possibilities as fiction, but it's ability to evoke response is the ELIZA program. ELIZA, also known as DOCTOR, was written in 1964 by Colby with Joseph Weizenbaum, a computer scientist at MIT. The program acts like a therapist asking leading questions then mirroring the user's answers. The following conversation is typical. ELIZA starts right out in the therapist role:

HOW DO YOU DO. PLEASE TELL ME YOUR PROBLEM.
I can't find my car keys.
TELL ME MORE ABOUT YOUR CAR KEYS.
I thought I left them on the hall table.
YOU LEFT THEM ON THE HALL TABLE?
I'm not absolutely sure.
WHAT MAKES YOU SO UNCERTAIN?
I forget.
CAN YOU THINK OF WHY YOU MIGHT FORGET YOUR CAR KEYS?

The topic can of course be much more personal, and is most successful on that level because ELIZA scans your input sentences for key words having to do with family, dreams, and emotional states. Her responses seem a little more responsive, a little less canned, if you stick to her domain. She is a passive conversationalist though, by the very structure of the program. Key words are her only method of cueing a response, so if she doesn't find a key word she has to use a stalling tactic until she gets one. ("PLEASE GO ON.")

Because the extent to which this system actually "understands" is limited to matching input words to words in memory her seeming intelligence depends heavily on clever schemes and our natural anthropomorphism. In the therapist role ELIZA is always a listener, never called upon to bring up subjects on her own. The context imposed by the therapist role is quite narrow, any subject the user brings up will immediately be drawn back to the user's own thoughts and feelings, which is engaging to a limited extent, and makes it unnecessary for Eliza to know anything of the world outside this interaction.

The design trick is to make everything she says general enough that she will never say anything that is too obviously wrong. It is easy to exhaust her bag of tricks though, unless you are a cooperative user, helping her reinforce the illusion of conversation by avoiding the things that trip her up. And it's surprising how cooperative people will be.

Weizenbaum, Colby's partner in creating ELIZA, was increasingly skeptical about representing this program as artificial intelligence. In his book **Computer Power and Human Reason** he talks about how people reacted to ELIZA (here, in one of her early

incarnations called DOCTOR) at MIT where the program was on the system and often used as a demonstration program:

> I was startled to see how quickly and how very deeply people conversing with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. Once my secretary, who had watched me work on the program for many months and therefore surely knew it to be merely a computer program, started conversing with it. After only a few interchanges with it she asked me to leave the room. Another time, I suggested I might rig the system so that I could examine all the conversations anyone had had with it, say, overnight. I was promptly bombarded with accusations that what I proposed amounted to spying on people's most intimate thoughts; clear evidence that people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms.

Actually it sounds to me as if people conversed with the computer in a *more* personal way than if it was human.

It is in fact amazing that such austere line drawings of intelligence can arouse in us such strong identification.

Weizenbaum and a few of his colleagues grew disenchanted with these programs, they were afraid that they were presenting people with a kind of counterfeit interaction. Dismayed that this token "intelligence" was accepted at face value as readily as it was by both the AI community and the public at large Weizenbaum wrote his book as an examination of Western culture's too-ready embrace of a rational/ mechanical picture of the human mind.

AI has had a mindset that reminds me of Renaissance enthusiasts of realist painting who at some point, the new power of three-point perspective gone to their heads, made claims that they would advance the art to the point of capturing actual reality.

An essay by philosopher Robert Sokolowski[3] lays out the development of artificial intelligence as just another step in the history of the articulation of meaning. This step

_____

[3]Natural and Artificial Intelligence", in the Winter 1988 issue of the quarterly **Daedalus.**

follows the conceptual jumps humans have made in going from the "brain word" (mental representation) to the spoken word, from the spoken word to the written symbol (first ideograms, then our more abstract alphabet) and now artificial intelligence taking it a step further, setting it in motion, making an abstract of thought.

Thinking of artificial intelligence not as an attempt to reproduce the mind per se, but as a step in the evolution of the symbol opens it up to being not an imitation but an auxiliary form, makes it more than computer science, puts it in the realm of the humanities.

Story

A couple of years ago my thoughts on all this distilled to the point that I could put them into story form, and I wrote a video screenplay. Dramatically titled "A New Face in Hell", it was about problems inherent in making minds. There were two main characters in the story, a computer and a young woman who was a graduate student in cognitive science. She was programming the computer, making a simulation of personality, trying to get all the right factors working to make it act human.

As progress is made her program, MARGO, does take on personality, and a familiar personality at that. Because much of the data she used for her model was taken from familiar sources, from her own experience, the programmer starts interacting with the model personality in ways similar to ways she interacted with the person she is most familiar with, her own mother. She rationalizes the use of her mother's life as a database by telling herself that she will only know if it is internally consistent if it is a model of someone she knows intimately. (She does exclude all history of herself as Margo's daughter from MARGO's database. It doesn't help, she exists tacitly in the computer's world as a daughter because the data is collected, inevitably, from a daughter's point of view.)

The situation of using such a subjectively loaded model is of course exaggerated, but with the intent of underscoring the impossibility of escaping our own experience, our own constant state of bias when it comes to thinking about thinking. MARGO does only what she programmed it to do. It is in fact very limited in that it's only "human" feature is speech, it can talk. But the familiar expressions and turns of phrase that the

computer uses, even in it's thoroughly synthetic voice, prove to evoke this woman's mother so powerfully for her that she often forgets these are words she herself has given the computer to say. There is a feedback problem of sorts, the programmer's own programming kicks in and she can no longer examine the situation because she can't get outside of it, can't avoid reacting to it, she is limited by invisible assumptions which have to be rooted out.

The story is a cautionary tale, I suppose, pointing to our vulnerabilities when it comes to seeking meaning, but I also meant it to prompt the question of whether artificial intelligence, the portrayal of thought, is necessarily an exclusively scientific pursuit.

After spending a great deal of time writing the screenplay and making preliminary attempts to produce it, it occurred to me that the most interesting aspect of the script was not particularly narrative. It did not need to be part of a story and would in fact be much more interesting as first hand rather than interpreted experience. I decided that making an installation in which people could work with MARGO, the computer character, would be more fun for everyone concerned.

So, like the programmer in my screenplay I set out to make a computer personality.

MARGO

Making Margo a parental persona was simply what I was drawn to do, but it was fortuitous choice in that, like the analyst personality of ELIZA, the parental role had many built in advantages. First, the fact that there is a body of experience in early memories that we all hold in common would make it possible for anyone, young or old, to have a sense of familiarity with the routine. The second advantage implicit in this role is that she is in charge. She is the bossy person always telling you what to do and only listening to you peripherally. This is admittedly an exaggerated characteristic in Margo, she is bossier than most. As a design strategy this is the direct opposite of a design like ELIZA's passive listener which is completely dependent on the listener to drive the interaction. Margo has an aggressive structure, she gets the first word, the last word, and if the user does not respond she asks them why not. This characteristic charging ahead with a topic makes it possible to maintain the illusion of conversation

even when her replies are less than responsive. It's in character for her to occasionally "ignore" you.

Using a computer I hoped some distance could be gained, a tension could be created by having these intimate things be said by a voice that is not human with an expression which is not a face, using an indirect messenger. In being a machine and obviously not a real emotional being Margo provides a neutral backdrop for something that can otherwise have no neutral presentation - we would not accept an actress saying Margo's lines as an interlocutor in the same way. Like text on a page, the things that Margo says are much more clearly words. It makes it possible to shift focus from the expressions as things that have emotional freight to things that are sounds made by a machine.

Background

The background in a painted portrait will often give some sense of the world that the subject inhabits. The video clip that plays as a prelude to each conversation with Margo is meant to serve an analogous function, to provide some context, to call up from memory the early world where she is a protagonist.

There are 32 different clips, 5 to 10 seconds long, pressed on a video laser-disc. Each of these clips is a typical childhood experience: a grown-up hand reaching back to take yours before crossing the street; a survey of the rubble on the floor of your room; watching your hand playing in the rush of air out the car window; staring down into the forbidden deep end of the pool. All of these are shot point-of-view style from a child's perspective, and Margo chimes in with a typical remark during each of these little memories. For instance during a clip where an ice cream cone dribbles down your arm Margo says, "Now don't you get that on your clean shirt!".

The video image then freezes on the screen and Margo launches into a conversation, (after the ice cream cone clip the conversation is about how important it is to maintain a neat appearance). At this point Margo begins directly addressing the user, and a set of possible responses appears on the computer screen. The response options range in tone from sweetly cooperative to openly belligerent. Using the computer mouse you select a reply, then Margo will say something back and the conversation continues until the

subject is changed by either the user or Margo herself. Another video clip will come up to introduce the next conversation.

Some of the video clips, rather than being prelude to a conversation with the user are the introductions to cautionary tales; warnings about the hazards of eating before swimming, reading with too little light, sitting too close to the TV. No user response choices will appear on the computer screen during these stories, instead an admonition to sit still and listen appears.

Also visible on the computer screen, behind the response choices, is a rectangle which constantly changes size, shape, and shade. As Margo's mood improves the rectangle gets larger and lighter in shade. Bad moods are indicated by a smaller and darker version of the rectangle. Like a facial expression it gives you a little extra information about the meaning of what is being said, and it gives you an immediate visual indication of the effect of your last reply. (A more detailed description of the expression follows under the Emotions heading.)

So the full installation includes not only Margo, but a video background, and, as a framing device Margo and the video monitor sit in a kitchen table environment. **Early Programming** is the title of the piece as a whole.

Early Programming is not so much a portrait of a single person as it is of a relationship. The parental role and its counterpart - in every parent a child. So, not a portrait of any one individual, but a portrait of an exchange.

Emotions

Margo is not essentially a meaning machine. Conversing with her is an exchange of emotional tokens, not ideas. The general course of interaction with her is determined by her mood which the user can effect to some extent, but not determine entirely. Margo's emotions, such as they are, issue from the emotional engine, a part of the main program.

The emotional engine works as follows:

Margo has an adjustable internal cycle which, independent of input, goes up and down over a period of time. All other emotional processing rides this cycle. This gives her a changeable basic mood which is affected by the user's responses but will continue to go up and down independent of input. When the user chooses a response the value attributed to that response comes in to the emotional engine and is first evaluated as to whether it's value is different from the last response you gave her. If this response has a value the opposite sign of the previous response it increments her stimulation level slightly. If it is the same sign as your last reply it neither increments or decreases stimulation, but the stimulation level falls slowly over time unless incremented.

In more anthropomorphic terms, Margo is expecting varied responses and unpredictable conversation. She gets bored if all your answers are always petulant, or even if they are always nice, well-behaved responses. She will tend to fall into her less energetic affects, though this may be offset by the base mood. If there is a good mix of positive and negative answers she will stay in the more energetic affects. To continue in the anthropomorphic vein, she cannot be immediately won over with a high scoring reply because she remembers your replies over time and their effect diminishes slowly, except in certain cases where a volley of high scoring replies can have the cumulative effect of a rush of good spirits. Similarly a barrage of negative scoring replies can lead to a burst of fury. After such an outburst, though, she will quickly return to equanimity.

There is no linear scale of emotions, one leading to another, so figuring out the spectrum of Margo's ups and downs was not a simple problem. Margo is fortunately not required to have a full range of emotions, she is only a slice of a full personality, only a role. I started with the notion that I could rate emotional affects by the amount of excitation or stimulation associated with each, so that hilarity and rage, a positive and negative state of high excitation are roughly equivalent but at opposite ends of the scale. Margo's particular scale goes from +4 to -4 .

The DECtalk voice synthesizer which generates Margo's speech allows you to adjust more than twenty different voice parameters making it possible to design a voice and then by adjusting certain variables in the voice to make it, to some degree, expressive of a particular emotional state. For instance increasing the breathiness makes it sound more intimate and friendly. When the pitch range is at maximum the voice can sound quite animated, and sped up, it sounds excited. If you decrease the speed and the pitch

range it can sound monotonous, depressed. Each of Margo's nine affects has different voice specifications which are given to the DECtalk each time she says something.

Following is a description of my interpretation of each emotional affect and the voice specifications that go with it:

Hilarity - In this mood Margo jokes, says silly things, makes light of situations, uses ridiculous and embarrassing expressions like "Yoo -Hoo!". In this mood she is part of the grown-up world which finds children inexplicable little darlings and makes fun of them sometimes in thoughtless ways. In a positive light, sometimes the things she says in this mood catapult you out of the deadly seriousness of childhood's point of view. She acts like a mother might act when she is with her grownup friends. Her voice has maximum pitch range, goes up and down quite a bit and is as expressive as the voice synthesizer allows.

Enthusiasm/Interest - In this mood Margo is full of optimistic expectations, advice, constructive criticism. She believes in you, and your problems can be solved if you just try. She wants you to live up to your potential and to make her proud, she is closer emotionally than in hilarity, she is identifying with you. Her voice is rapid and bright.

Warmth - Here Margo is a friend and ally who is truly sympathetic even when critical. She also has a sense of humor which includes rather than excludes you. In this mood she is a grownup in the reassuring sense that she has a broader perspective than you do. If she says things will be okay you can believe it. This is definitely the nicest Margo, the closest emotionally and the most persistently rational. She speaks slowly and quietly.

Neutral - Margo's neutral persona most baldly states the topic and attitude of each conversation. It corresponds to the situation where you're not sure what's going on, maybe nothing. It's the state where Margo is least maternal and most matter of fact. Her vocal characteristics are all set at medium.

Glum - Here Margo descends into a less felicitous mood. Her voice is flat and dull. She complains, she tries to inflict guilt and starts many sentences with, "I don't suppose.."

She has a self-pitying attitude. In this mood she is distant and her expectations are not high.

Sarcastic/Irritated - In this affect Margo turns things around, saying the opposite of what she means, or else she speaks with a world weary irritation. She is not treating you like a child, really, she doesn't expect you to understand what she's saying. The things she says can be funny and tension relieving, but sometimes they have an edge to them. Her voice is slightly monotonous and she speaks slowly with long pauses.

Huffy - the huffy affect is cold and short and angry, always in a direct way, no sneaky insinuations. She is obviously mad at you. She tends to imperiousness and calls rank on you. ("You will do it because I say so!") In this mood she has expectations of you and you have let them down. Her anger is an expression of closeness temporarily withheld. Her voice is sharp and she speaks shortly.

Dismayed - In this mood Margo is the adult who is no more in control than you are, who is in fact depending on you to some degree, and is alarmed and upset when you are unhappy, feeling that you're letting her down. In this affect you get blamed for upsetting her, not for what you've done. It's also the affect where she is most concerned with what other people think. She sounds a little hysterical, her voice is high-pitched and rapid.

Raving - Is not as bad as it sounds. The longest sentences and the most overwrought language, but she's not as seriously angry as when she's huffy. It's the fullest version of all the clichés of the overworked underappreciated self-sacrificing parent. The fullblown threats of violence and abandonment are part of this mood, but they're rhetorical, not venomous. It is her opportunity to use lots of colorful speech, and sometimes its the most straightforward presentation of the situation. She speaks expressively but rapidly.

The shaded rectangle, or expression, is the visual manifestation of the emotional engine. The rectangle is meant to convey, over time, a sense of her emotional ebb and flow. The width of the rectangle is determined by the level of stimulation, and the height is determined by her mood, which is an expression of the cumulative affect of the last few answers. The expression is redrawn every thirty seconds. The gray level and texture of

the rectangle is determined by her current affect, the darker, denser ones being the angriest.

While working on the program I was encouraged to make the rectangle into a drawing of a real face with different expressions for different moods. The idea being that the face is the ultimate in readability for human face-readers. I am opposed to doing this for the simple reason that I think that it would take away from her machine-ness by making her into an animated character. Remaining an abstraction is important in order to be able to maintain a shiftable focus, for the viewer to be able to see both the illusion and the apparatus.

I would however like to make the expression more emotionally evocative, perhaps by using color, not in a directly symbolic way (red is anger, blue is peaceful , etc.) but in combinations that are more and less harmonious, vivid or varied. Figuring out which combinations will evoke different emotions is a project in itself, which I look forward to doing.

Conversations

> "I then began again to think about the bottom nature of people, I began to get enormously interested in hearing how everybody said the same thing over and over again with infinite variations but over and over again until finally if you listened with great intensity you could hear it rise and fall and tell all that there was inside them, not so much by the actual words they said, or the thoughts they had but by the movement of their thoughts and words endlessly the same and endlessly different.
>   -Gertrude Stein, **"The Gradual Making of the Making of Americans"**

Most of the phrases and expressions Margo uses are things I remember from my mother, my grandmother, or other kid's mothers, and the rest are things other people remembered hearing. The conversations are like folksongs, I didn't write them so much as arrange them.

I organized all these sayings around events that happen over and over in childhood. Leaving the house, going to sleep at night, eating at the table.

Margo's end of the conversation is written out in steps each consisting of a phrase that relates to the step before it and the step after it, and carries forward some particular topic. Every step of a conversation is written in each of the nine different affects, so that the same thing is said with enthusiastic interest, sarcasm, warmth , etc.

Writing the conversations is a little like doing a crossword puzzle - things have to work in more than one dimension.

FRESH AIR + EXERCISE			
HILARITY	I CAN'T BELIEVE MY EYES! WHAT ARE YOU DOING INSIDE ON A DAY LIKE THIS?	WHAT'S THE MATTER, FLOPSY? AREN'T YOU FEELING WELL? DON'T YOU WANT TO GO PLAY?	ALL THE OTHER CHILDREN IN THE NEIGHBORHOOD ARE OUT THERE HAVING A WONDERFUL TIME! NOW COME ON, LET'S GO...	NOW LISTEN HERE, SWEETIE, I HAVE THINGS TO DO IN HERE AND YOU DON'T WANT TO JUST HANG AROUND + WATCH. GO ON - SCOOT, OUT YOU GO!
ENTHUSIASM	WHAT A DAY! YOU DON'T WANT TO WASTE IT IN THE HOUSE!	YOU HAVE TO TAKE ADVANTAGE OF WEATHER LIKE THIS - GO FOR A WALK IN THE SUNSHINE!	CALL ONE OF YOUR FRIENDS, I BET YOU CAN FIND SOMEONE TO PLAY WITH.	IT'S GOING TO BE A BUSY DAY! I HAVE A LOT TO DO HERE, AND YOU HAVE LOTS TO DO OUT THERE. OKAY? OKAY.
WARMTH	OH BOY, WHAT A DAY IT IS OUT THERE TODAY. WANT TO GO OUTSIDE?	BETTER GET OUT IN THE SUNSHINE WHILE IT LASTS. YOU'LL FEEL FINE ONCE YOU GET GOING.	GO GET ON YOUR SHOES AND JUST GO OUT AS FAR AS THE MAILBOX AND SEE WHAT'S GOING ON. OKAY?	I'D GO WITH YOU - IT'S SO NICE OUT THERE, BUT I'VE GOT TO PAY BILLS + VACUUM
NEUTRAL	IT'S A BEAUTIFUL DAY, WHY DON'T YOU GO OUTSIDE AND PLAY?	YOU'VE BEEN INSIDE ALL WEEK, NOW GO OUT AND RUN AROUND.	FIND SOME OF YOUR FRIENDS. GO RIDE YOUR BIKE. WALK THE DOG.	WELL IF YOU'RE GOING TO STAY IN THE HOUSE I WANT YOU TO HELP ME.
GLUM	FOR HEAVEN'S SAKE, WHAT ARE YOU DOING INSIDE ON A DAY LIKE THIS?	ARE YOU TRYING TO TURN INTO A VEGETABLE?	YOU'D RATHER SIT IN THE DARK AND WATCH TV THAN LEAD A NORMAL LIFE.	I DON'T KNOW WHAT'S THE MATTER WITH YOU. I JUST DON'T UNDERSTAND IT.
SARCASTIC	WELL THAT'S A GOOD THING TO DO ON A SUNNY DAY. LET'S ALL LIE ON THE LIVING ROOM FLOOR.	THERE'S NOTHING AS INVIGORATING AS LYING AROUND IN YOUR PAJAMAS, IS THERE...	I BET YOUR BUDDIES AREN'T SITTING AROUND INSIDE, MOPING.	IF YOU REALLY WANT TO, THOUGH, YOU CAN HELP ME WAX THE FLOOR.
HUFFY	I WISH YOU WOULD GO OUTSIDE AND PLAY.	IT DRIVES ME CRAZY TO SEE YOU LOUNGING AROUND IN HERE ON SUCH A NICE DAY.	IF YOU WEREN'T SO LAZY YOU COULD BE OUT HAVING FUN.	IT'S EITHER GO PLAY OR STAY HERE AND MAKE BEDS, UNDERSTAND?
DISMAYED	WHAT ARE YOU DOING INSIDE? I THOUGHT YOU'D BE OUTSIDE, PLAYING.	ARE YOU SICK? SO YOU WANT TO GO BACK TO BED? MAYBE YOU SHOULD BE IN BED...	ALL YOUR FRIENDS ARE OUTSIDE!	I'VE GOT ALL THIS HOUSEWORK TO DO AND NOW I HAVE TO ENTERTAIN YOU, TOO?
RAVING	WHAT'S THIS? HAVE YOU LOOKED OUT THE WINDOW? IT'S A BEAUTIFUL SUNNY DAY AND YOU'RE MISSING IT!	YOU SHOULD BE ASHAMED OF YOURSELF MOPING AROUND IN HERE ON A DAY LIKE THIS. GO GET SOME EXERCISE! IT WON'T KILL YOU!	GO TO THE PARK! GO RIDE YOUR BIKE! PLAY BALL, WALK AROUND THE BLOCK, I DON'T KNOW... JUST GET OFF YOUR DUFF!	I DON'T WANT YOU IN HERE. I HAVE ENOUGH TO DO TODAY WITHOUT LISTENING TO YOU COMPLAIN! OUT!

Fig. 1.
Example of a conversation diagram (from the "Fresh Air and Exercise" scene) with steps in the conversation along the horizontal axis, and emotional states along the vertical axis.

The tricky thing is that though Margo may enter a conversation in a huffy affect, given a well-behaved response her mood may turn to only being glum, or given a contrary response might go into a rage. Her mood can change from one conversational step to another, so all conversational steps have to be modular, interchangeable, and still make

17

sense, though the "sense" Margo makes is largely interpreted. Her conversation is largely cliché, which is actually a form that leaves lots of room for attribution.

Repeatedly writing a conversation in different moods is an interesting exercise in uncovering what is actually going on in that particular exchange of words. Examined in this light many of the conversations which had seemed funny or quaint show less frivolous aspects, the bones of a power relationship. A preponderance of the things I remember as stock phrases have to do with exerting control or resisting, denying the right of the other to control. But the gist isn't so much domestic imperialism as it is social subroutine.

Having the user's response choices on-screen, on buttons, (they're called buttons as a graphic analogy to pushbuttons) was not my original intent. I had intended that people make up and type in their own responses. For the first demonstration of the program, though, I put buttons on-screen because MARGO was not yet capable of taking keyboard input. As it turned out the buttons did an important thing that keyboard input does not, that is they introduce the role you play with her. The response choices are clearly characteristic of child/parent repartee. There were numerous suggestions that I keep the buttons, and make more of them (the first version had only four buttons). Using a limited number of responses had also the great advantage of simplifying the programming by sidestepping the issue of word and phrase recognition by the computer.

The response buttons are each scored for their agreeableness or disagreeableness. A +5 is the best, something on the order of saying you will immediately go clean your room, you'd like more brussel sprouts please, you're very sorry, you'll never do it again. A -5 will prompt Margo to ask what has gotten into you, why are you so crabby and insubordinate? A -6 she interprets as a contradiction. This usually brings a harsh response and also contributes richly to a bad mood. A -7 is reserved for the use of obscenities and bad language, which Margo abhors. This more than anything will make her wrathful, and she will scold you.

Rating responses, giving them scores, was by no means a straightforward procedure and I usually followed a sort of intuition about it rather than making any formal rules. There were some things that would obviously make her angry or pleased, but there

were areas where more obscure emotions came into play. Would she be proud of you for standing up to her, or angry at your insolence? Would she be glad you wanted to stay home with her or distressed at your lack of independence? The decisions about how to score different responses definitely lend idiosyncrasy to Margo's character. This is an obvious area where my own character is mapped onto hers.

Technical

Margo is programmed in HyperCard, a programming environment that runs on all Macintosh computers. HyperCard serves as an organizational framework for storing information and also allows you to write programs which will determine how the information you've stored is accessed. The overall analogy is that of putting information (text, numbers, pictures. musical notation, etc.) onto cards in a card catalog. The cards are grouped in stacks. Each stack is named and has its own rules of access and system of linkages from one card to another, perhaps linking to other stacks.

The Margo program consists of a central controlling stack and a group of conversation stacks which each hold all the possible phrases for one of her conversations. The conversation stacks are two-dimensional arrays in which each phrase is addressed by what time step it is in the conversation and its affect rating. The conversation stacks are called up by a central stack called the conversation manager. The conversation manager (Conman for short) has two main functions. First, to keep track of the emotional engine, feeding the input scores from the buttons to it and keeping track of the output giving a constant readout on changes of mood and stimulation level. From that readout Conman generalizes an affect and draws the expression. Conman then tells the DecTalk voice synthesizer what vocal characteristics to use for the given affect. Second, Conman is in charge of choosing what Margo says next. She will move on to the next step in a conversation unless the response she receives is an interrupt. Direct contradictions, bad language, and very crabby responses are interrupts and will send Conman directly to conversation stacks which hold responses specifically for that sort of thing. Conman also keeps track of which conversations Margo has had lately, and what is on-screen - which buttons, which video clips.
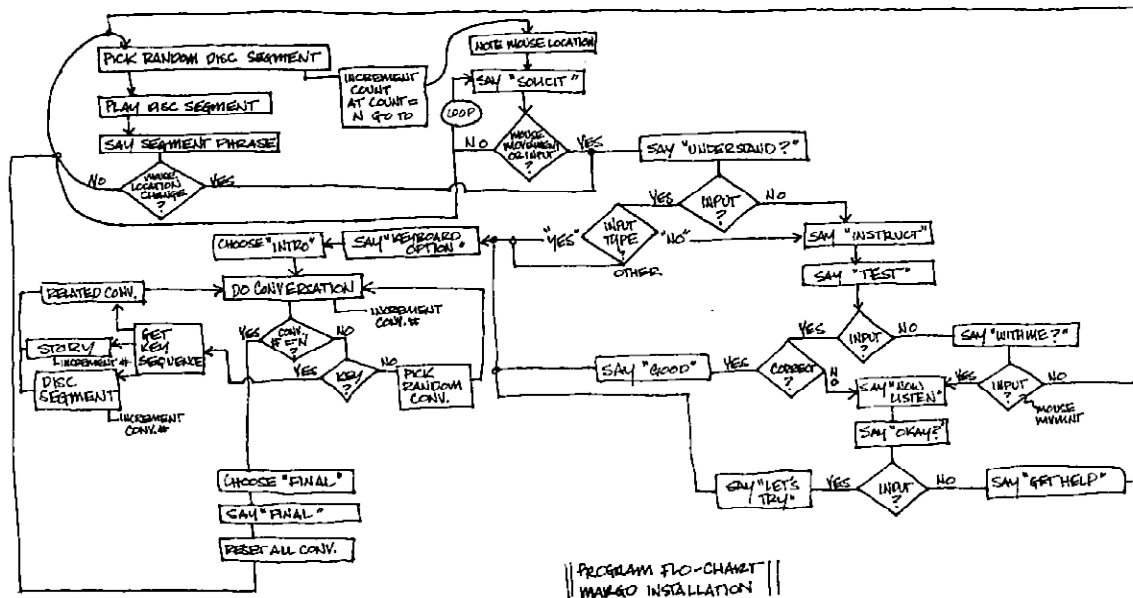
Fig. 2.
Flow chart for the Margo program.

When the mouse has not moved in some time Margo will go first into a "Cat got your tongue?" mode. She will ask several times what's the matter, why aren't you talking to me, with various degrees of heat depending on her mood. After she's made a certain number of these inquiries she will give up and start trying to get someone else to sit down and talk, by showing a few of the video clips and then pleading with anyone who might be in earshot to come talk with her.

The teaching loop is designed to have Margo teach people how to use her in a way which is in character. The teaching conversation is also mood driven but in the interest of not scaring anyone away she remains in the four most neutral moods. She goes into the teaching loop when the mouse is first moved after a pause. She immediately asks the user, "Do you know what you're doing with that?" There are yes and no buttons on-screen and if she gets a yes she then goes on to the rest of the program. If she gets a no answer or just mouse movement she will go on to explain mice and buttons to the novice user, testing them and eventually leading them into a conversation. If there is no movement at all she tries to persuade the erstwhile user to at least move the mouse. If this does not work she tells the user who in the gallery will help them, and she goes back to trying to find a new user.

Further

Currently Margo is randomly driven, that is in choosing a conversation her only strategy is to choose one that she hasn't had lately. I would like to change this so that the conversation is determined more by the interaction in the previous conversation(s). Her affect level carries over from conversation to conversation but it would be nice if the topics were somehow more emotionally related as well. For instance, if responses are disrespectful her next conversation could be about insolence. If they are impatient she could go on to tell you about the virtues of patience, if you are particularly agreeable she might praise you or reward you with a story. To do this is a matter of making Margo able to discern a higher level of pattern in the user's responses, and able to proceed according to what she finds.

At one point I considered a strategy in which Margo starts out treating you more like an adult, with conversations asking you whether you've gotten your car insurance yet, are you ever going to settle down, where did you get that awful haircut, etc. She would then regress to teenage conversations,   (You're going to kill yourself driving that way, you're not going out of the house dressed like that!) before finally going back to the childhood conversations. It would be interesting to carry these two aspects through to give a sense of the rise and fall of parental power, the shift in the relationship. But I think it might call for a much more elaborate design. I'm not sure that the video clips leading to conversations format would work for that. Visual entrees for adult conversations might not be universal enough.

Finally

My interest in was initially in how we relate to an object as alive, what there is about it, what there is about us, that promotes these hazy notions of a thing as an entity with a fate of its own. In the early video tapes, I was exaggerating and commenting on this tendency. Then, with **Early Programming**, I decided to make an object which attracted that kind of attribution but which also allowed the user to observe it as a thing, a machine. As you might expect, going from being the subject of the phenomenon to being the one attempting to provide this peculiar experience I noticed a thing or two, and my point of view changed.

At the onset I felt some suspicion, I felt critical of this tendency, this anthropomorphizing of things. Like Weizenbaum perhaps, I felt that we were being fooled into these assumptions, somehow grabbed in a tender spot and led down the garden path. Or that maybe we just had a perverse penchant for fooling ourselves to no good end. Much of my initial interest in this was a sort of cub reporter zeal for laying it bare.

What gradually occurred to me was that perhaps this animation is not necessarily a lazy or dangerous substitute for understanding, but one of the first steps in knowing, a sort of primary grappling with abstraction. That faced with the unfamiliar the first mental tentacle outstretched may be a sort of emotional scout looking for a way to feel about this thing. And that whatever impressions that first contact sends back may be especially potent in shaping whatever understanding comes later.

The other realization, which sunk in with some struggle, was that programming a computer is an act of abstraction, a matter of creating an artifact which not only performs some function but also expresses its creator, for better or worse. Another medium, another craft. I look forward to it's proliferation.

Bibliography

Athanasiou, Tom   "Mind Games" in *The East Bay Express*, Vol. 7 No. 9, Dec.7, 1984

Bateson, Gregory   *Mind and Nature* New York, E.P. Dutton,1979

Boden, Margaret   *Artificial Intelligence and Natural Man* , New York, Basic Books, Inc., 1977

Cohen, Harold   "What is an Image?" in *Proceedings of the Sixth International Joint Conference on Artificial Intelligence,* Tokyo, 1979

Colby, Kenneth M.   *Computer Models of Thought and Language* eds. with Roger C. Schank, San Francisco, W.H. Freeman and Co., 1973

Dreyfus, Hubert   *What Computers Can't Do: A Critique of Artificial Reason*, New York, Harper & Row, 1972

Dennett, Daniel C.   "When Philosophers Encounter Artificial Intelligence" in *Daedalus,* Vol. 117, No.1, Winter 1988, American Academy of Arts and Sciences

Gonick, Larry   *The Cartoon Guide to Computer Science,* New York, Harper & Row, 1983

Haugland, John   *Mind Design: Philosophy, Psychology, Artificial Intelligence,* editor, Cambridge, MIT Press, 1981

Hillis, W. Daniel   "Intelligence as an Emergent Behavior" in *Daedalus ,* Vol.!117, No.1, Winter 1988, American Academy of Arts and Sciences

Jackson, Phillip C, Jr   *Introduction to Artificial Intelligence* ; second edition, New York, Dover Publications, 1985

Johnson, George   *Machinery of the Mind,* New York, Random House, 1986

Miller, Jonathan   *States of Mind*, New York, Pantheon Books, 1983

McCorduck, Pamela   *Machines Who Think,* San Francisco, W.H. Freeman, 1979

Newman, James R.   *The World of Mathematics, Vol.4,* New York, Simon & Schuster,1956

Minsky, Marvin   *The Society of Mind*, New York, Simon & Schuster, 1985

Plutchik, Robert   *The Emotions: Facts, Theories, and a New Model*, New York, Random House, 1962

Schank, Roger C.   *Computer Models of Thought and Language,* editor, with K.M. Colby San Francisco, W.H. Freeman, 1973

Searle, John   *Minds, Brains, and Science*, Cambridge, Harvard University Press, 1984

Sokolowski, Robert   *Natural and Artificial Intelligence,* in *Daedalus ,* Vol. 117, No. 1, Winter 1988, American Academy of Arts and Sciences

Turing, Alan M.   "Can a Machine Think?" in *The World of Mathematics, Vol.4,* ed. James R. Newman, New York, Simon & Schuster, 1956

Turkle, Sherry   *The Second Self; Computers and the Human Spirit*, New York, Simon & Schuster, 1984

Weizenbaum, Joseph   *Computer Power and Human Reason; From Judgement to Calculation,* New York, W.H. Freeman and Co., 1976

Winograd, Terry   *Understanding Computers and Cognition*, with Fernando Flores, Addison & Wesley, 1985

Winston, Patrick   *Artificial Intelligence; an MIT Perspective,* with R.H.!Brown, editors, two volumes, Cambridge, MIT!Press, 1979