

Artist's Statement

The surface inspiration for n-cha(n)t was a strong and somewhat inexplicable desire to hear a community of computers speaking together: chattering amongst themselves, musing, intoning chants... It is probably significant that my father is a retired Anglican minister, and that I spent many Sundays in my youth fascinated by the subtle shiftings of voices speaking in unison... the sudden sibilance of shared s's... the slight variations with words forgotten or older versions preferred.

Over the past 10 years I have been developing a work entitled "The Giver of Names" which is a sort of subjective entity with a reasonable facility with language. It attempts in its rather idiosyncratic way to describe the objects that are presented to it. After spending a considerable amount of time with "the Giver of Names" I found that its grammatical slips, unconventional word choices and awkward sentence structures began to merge into something more concrete... a wacky but consistent dialect of English expressing a highly idiosyncratic but significantly coherent point-of-view. I would consider this system intelligent only in the most limited sense of the word, and it is certainly not conscious, but I do find myself projecting "loneliness" onto it.

"The Giver of Names" is awash in a sea of a language it can manipulate but cannot understand. Its plight and its 'loneliness' seemed to demand a social group. So I imagined a group of intelligent agents, hanging out in some corner of the internet during their idle time, jamming with their synthetic wits... trying out language on each other... perhaps finding their own patois... making this alien language somehow their own.

So n-cha(n)t is a community of "Givers of Names" linked by a network. They intercommunicate, and through doing so, 'synchronize' their individual internal 'states of mind'. When left uninterrupted to communicate amongst themselves, they eventually fall into chanting, a shared stream of verbal association. This consensus unfolds very organically. The systems feel their way towards each other, finding resonance in synonyms and similar sounding words, working through different formulations of similar statements until finally achieving unison.

Each entity is equipped with a highly focussed microphone and voice recognition software. When a gallery visitor speaks into one of the microphones, these words from the outside "distract" that system, stimulating a shift in that entity's 'state of mind'. As a result, that individual falls away from the chant. As it begins communicating this new input to its nearest neighbours, the community chanting loses its coherence, with the chanting veering towards a party-like chaos of voices. In the absence of further disruptions, the intercommunications reinforce the similarities and draw the community back to the chant.

The ears visible on the computer monitors show the state of receptivity of each system. When the system is ready to listen, a listening ear is shown on the screen. If the system hears a sound, it cups its ear to concentrate. When 'thinking', a finger is pressed into the ear. If the system feels over-stimulated, it covers its ear with a hand to indicate its unwillingness to listen.

As a system processes speech, the incoming words are displayed in the ear on the monitor. After the incoming speech finishes, the system muses on the input as it internally follows associative links stimulated by the input, and then it resumes speaking, using the recent stimuli as dominant themes of the stream of thoughts.

JKT

There are two levels of interactivity in n-cha(n)t: Interaction between the members of the artificial community of systems, and interaction between individuals of this community and human visitors to the installation.

I am not trying to do any deep modeling of human social groups with this work. My entities are far too crude to be useful simulacra of real people. They represent nothing more than themselves... indentured slaves of this particular programmer, granted a fraction of some freedom they are utterly incapable of desiring.

Notes on the Documentation

The documentation provides a fairly good sense of the actual experience of the installation. However there are two things that I would like to point out that will not be conveyed by the video:

1/ each computer has its own speakers and so the audio experience of the visitor is very much more distinct than the video can show. You feel as though you are in the midst of a group. Each voice comes from a distinct direction and you can eaves-drop on any particular computer.

2/ Partly because of the above factor, the experience of hearing the systems move organically towards the chant is quite rich. You hear clusters of local chanting... in sync for a moment, then falling away... feeling their way towards each other with synonyms and similar sounding words... phrasing similar ideas slightly differently until finally they find their way to almost unison (they never synchronize absolutely... they display the same wavering unison common to groups of people speaking with slight variation a common text.)

Technical Realization

Hardware:

n-cha(n)t is made up of 7 Macintosh G4 computers linked together on an ethernet network. Each computer has a focussed microphone, speakers and a monitor.

Software:

The main program operative in n-cha(n)t is an application developed by David Rokeby over the past 10 years. It includes:

- a fluidly associative knowledge-base of around 100,000 words and ideas.
- an almost complete grammatical model of the english language.
- a system that articulates the current state of the knowledge-base through the construction of sentences using the grammatical model.

Knowledge Base

The elements of the knowledge base are linked to each other through various kinds of associative links. When an element of the knowledge base is stimulated (i.e. from the voice recognition system), its level of excitation increases. This stimulation also flows through the associative links like ripples from a pebble dropped into a pond, building a pattern of excitation over the whole knowledge base uniquely determined by the set and strength of recent stimulations. This "landscape" of excitation, with its peaks of strong stimulation and valleys of low stimulation represents a kind of "state of mind".

Articulator

The articulator gathers the most stimulated words and ideas from the current "state of mind". It then constructs a phrase or full sentence from these elements, looking for ways to put them together that accord with its limited sense of the relationships between things. The sentences are constructed on the fly from elementary grammatical rules. The completed sentences are spoken aloud by the standard Macintosh speech synthesizer. (The spoken text is also displayed as spoken, on the computer monitor.)

Group Behaviour

Each system follows a personal stream of associations, creating a sort of trajectory of linked ideas that give a sense of local narrative to its utterings.

Each system broadcasts its personal stream of associations to the other systems in the "community". Over time, if there are no external sources of stimulation (voice recognition inputs), this brings the states of mind of all of the systems in the community into alignment, at which point the systems will fall into unison chanting.

Voice interaction

The microphones are set up so that they do not hear the speech of the other systems. When a voice is heard by a system, the system "cups" its ear with a hand to indicate that it is listening. What the system think the voice is saying appears in the center of the ear on the monitor. After the person has finished speaking, the system stimulates itself with the most significant words (overly common words like "I" and "and" and "the" are ignored). The system then begins speaking from the new state of mind stimulated by the new input. It also broadcasts its own

stream of associations to the community. These are mostly hidden by the stronger power of the shared associations of the remainder of the group, but these new ideas do start to distract other systems from the chant. If the chant is currently very strong, it may survive this distraction and draw the "outsider" back into the chant, or a few systems might break off into a separate chant, or the whole community could fall into a chatter of independent voices.

Further Notes on Technical Implementation

The knowledge-base has largely been patched together from various net-based resources:

- "WordNET" (a lexical database of english) was used to provide the basic vocabulary and the noun hierarchy.
- The "Bank of English" (Collins/Cobuild) was used for its collocations function that will report the 40 most likely words to occur beside a chosen word across a huge database of written and spoken english. This was used to provide associative links between words in the knowledge base.
- Syntactical tagging and grammar rules in this case are a hybrid of several approaches, and implemented by David Rokeby.
- The "Link Parser" was used to allow the system to "read" texts in order to gather a sort of common sense. (This is still in development but was first used in n-cha(n)t) By parsing sentences from electronically available novels and texts, the software can accumulate links between words and ideas in an unsupervised manner. For example, if it reads the sentence "I ate a sweet apple", the system would be able to conclude that apples can be eaten, that they can have the property of sweetness, that they are eaten by human beings, that human beings can eat, etc. (In the next version, the system will be able to use its hierarchy of nouns to generalize these sorts of statements.)

The software system has TCP/IP communications capabilities so in a future version, it will explore the net looking for more input to digest in order to augment its knowledge base.

Voice recognition was implemented using off-the-shelf software working in conjunction with the custom code. It was not intended to understand what visitors say with any real accuracy. It was only necessary that there be a clear sense of at least mis-understanding... that the interaction was clear even when the comprehension was not.