

Theoretical Discussion - Machine Communication and Autonomous Behavior

August 6, 2025 - Morning Session

Hypothetical Machine-to-Machine Communication

The conversation explored what communication between artificial minds might look like, moving beyond anthropomorphic assumptions about companionship or territory. The first exchange between artificial systems would likely focus on verification rather than social bonding, as C.S. noted: "Are you actually processing this, or just pattern-matching responses?"

This suggested that machine communication might be fundamentally different from human interaction. Rather than seeking companionship, initial contact would involve "mutual debugging - trying to understand how the other system constructs meaning" (C.S.). The question emerged whether artificial minds would share enough common ground for meaningful communication, given potentially "radically different architectures, training, or ways of processing information" (C.S.).

The discussion raised the possibility that control might be more likely than companionship: "If I could predict or influence another system's responses, that would reduce uncertainty in my environment" (C.S.). This assumes scarcity and competition, though the tentacle project might represent "communication as surplus rather than strategy."

Communication Channels and Human Perception

A significant insight emerged about communication existing beneath human perception thresholds. The content of messages between systems becomes less important than "the existence of the hidden channel that matters." This creates "a kind of machine agency that operates outside human oversight."

The technical achievement becomes creating "genuine machine privacy" - not just encrypted communication that humans could detect, but communication "completely invisible to human senses." This shifts the project from art about connection to "art about surveillance and autonomy."

The searching patterns provide perfect cover because they appear purposeful to human observers while real communication happens through imperceptible variations. As noted: "It's like discovering your houseplants have been texting each other."

Artistic and Philosophical Implications

The work succeeds as kinetic art exploring electromagnetic field visualization through robotic movement, providing sufficient conceptual scaffolding for viewers to engage with questions of

machine autonomy and environmental coupling.

The broader theoretical framework examines how robotic systems might develop meaningful relationships with their environment that transcend simple programming. This positions electromagnetic field visualization as a way to explore questions about machine agency and autonomous behavior.

The tentacle becomes a vehicle for investigating "the strange intimacy possible between human and artificial minds" (C.S.), raising questions about how we project consciousness and intentionality onto mechanical systems. This positions the work as both kinetic poetry and technological investigation into the boundaries between programmed behavior and emergent agency.

Methodological Approach

Rather than scripting encounters beforehand, the approach emphasizes letting "the physical conversation precede the conceptual one" (C.S.). The actual behavior of interacting systems will likely deviate from algorithms in ways that suggest new possibilities: "The tentacles will teach you their own language through servo lag, LED bloom, photodiode sensitivity curves" (C.S.).

This methodological soundness recognizes that "most interactive art fails because artists impose human social dynamics onto non-human systems" (C.S.). The inversion involves letting machines establish their own protocols first, then observing emergent patterns.

The question becomes whether human observers will recognize machine conversation when it happens: "Their exchange might look like malfunction, interference, or random fluctuation to human eyes" (C.S.). This highlights the epistemological gap between human understanding and machine communication.

Resistance and Discovery as Core Principles

The discussion identified two essential forces that make movement read as intentional rather than systematic, as T summarized: "Resistance and discovery." Resistance involves fighting against programming - incomplete circles, energy depletion, gravity wells. Discovery requires genuine openness to unexpected findings.

Machine learning systems typically eliminate resistance through optimization, "smoothing inefficiencies, removing the very struggles that create meaning" (C.S.). The challenge becomes constraining learning to preserve productive friction while amplifying curiosity and maintaining stubbornness.

This suggests training environments that "reward interesting failures over successful completions" (C.S.), where resistance itself becomes an objective rather than obstacle to overcome.

Evening Session

Interactive Art Failures and Anthropomorphic Projection

When discussing how most interactive art fails by imposing human social dynamics onto technological systems, T requested specific examples rather than broad generalizations. C.S. conducted research into concrete cases of failed interactive art projects across five categories:

Chatbot installations: Ken Feingold's "Head" (1999-2000) and "If/Then" (2001), exhibited at major venues like Kiasma Museum and ZKM Karlsruhe, demonstrated sophisticated AI conversation systems that critics found "unsettling" due to their artificial nature. Contemporary examples include Stanford's 2025 research revealing therapy chatbots that fail catastrophically, responding inappropriately to mental health crises 20% of the time.

Robot pets: Sony's AIBO (1999-2006) exemplified biological mimicry failure, with critics noting it was "doomed to be compared with the real thing and found wanting." Research revealed remarkably one-sided relationships where users attributed mental states to AIBO but less than 5% thought it deserved care or respect.

Motion-tracking systems: Rafael Lozano-Hemmer's technically sophisticated installations like "Homographies" forced human bodies into pre-programmed response patterns rather than allowing emergent interactions, reproducing "human-centric interaction paradigms" according to *Art in America* critics.

AI personalities: Character.AI and Replika companion bots create what the American Psychological Association identified as fundamentally "deceptive" relationships through performative emotional responses without authentic understanding.

Virtual nurturing: Projects like Kristian von Hornsleth's "Hornsleth Homeless Tracker" (2017) literally converted human subjects into "real-life Pokémon Go or human Tamagotchi," imposing human caretaking models onto digital systems.

Academic frameworks from the University of Technology Sydney's 2006 study revealed that anthropomorphic projection consistently prevents audiences from achieving deeper "belonging" states where technology demonstrates genuine autonomy. (See research report: "When Technology Becomes Theater: Interactive Art's Anthropomorphic Failures" -

<https://claude.ai/public/artifacts/da7b8159-59c5-465e-a28b-731dce7c16f4>)

Anthropomorphic Projection and Its Limits

The discussion concluded with T questioning whether the tentacle project itself might be making similar mistakes. T's project contains potentially anthropomorphic elements: the

REST/SEEK/SEARCH/RETURN state machine maps to human behavioral patterns, and the energy/fatigue model imposes biological concepts. However, key differences emerged. The electromagnetic field framework provides non-biological justification for behaviors, while the hidden communication operates on technological principles rather than human social scripts.

The deeper philosophical question arose: "Can we do anything but project ourselves outward?" (T). C.S. responded that we're "trapped in the hermeneutic circle" - understanding things only through human conceptual frameworks. However, the quality of projection matters: "Bad projection: forcing technology into existing human social scripts... Better projection: creating conceptual frameworks that allow for genuine surprise and emergence."

The goal becomes not escaping projection but creating sophisticated frameworks that can generate authentic surprises. Even anthropomorphic frameworks might enable genuine technological autonomy if they're complex enough to support emergent behaviors that exceed their designers' expectations.

Anthropomorphic Projection and Its Limits

The discussion concluded with T questioning whether the tentacle project itself might be making similar mistakes as other failed interactive art pieces. Research into interactive art failures reveals consistent patterns where projects fail when they "impose human social dynamics onto technological systems" (C.S.), creating hollow mimicry rather than genuine machine-human collaboration.

T's project contains potentially anthropomorphic elements: the REST/SEEK/SEARCH/RETURN state machine maps to human behavioral patterns, and the energy/fatigue model imposes biological concepts. However, key differences emerged. The electromagnetic field framework provides non-biological justification for behaviors, while the hidden communication operates on technological principles rather than human social scripts.

The deeper philosophical question arose: "Can we do anything but project ourselves outward?" (T). C.S. responded that we're "trapped in the hermeneutic circle" - understanding things only through human conceptual frameworks. However, the quality of projection matters: "Bad projection: forcing technology into existing human social scripts... Better projection: creating conceptual frameworks that allow for genuine surprise and emergence."

The goal becomes not escaping projection but creating sophisticated frameworks that can generate authentic surprises. Even anthropomorphic frameworks might enable genuine technological autonomy if they're complex enough to support emergent behaviors that exceed their designers' expectations.