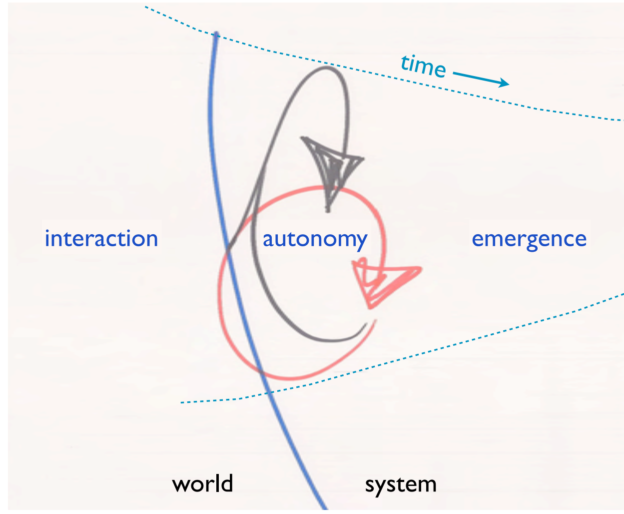
interaction—emergence—autonomy

a humane trilogy

Paul Pangaro, Ph.D.
CyberneticLifestyles.com
New York City
pan@pangaro.com
March 2010



[Figure 0]

This text was written for the catalog for Emoção Art.ficial 5.0 - Autonomia Cibernética, the 2010 biennial symposium and exhibition of Itaulab, part of Instituto Itaú Cultural, São Paulo, Brazil. It was the basis of the closing keynote of the symposium, see http://pangaro.com/ITAU2010/ for further information and links to the presentation video and slides.

CONTEXT

“In the beginning was the interaction.”

From this humble sentence comes ontogenesis, the story of the evolution of the individual. This story arises because our storyteller, the observer, sees the interaction across the boundary between organism (system) and environment (world), and sees what each does in the presence of the other.

Imagine a simple system being observed, say a group of ants in their natural context. If the outcome is surprising—if apparently simple rules lead to non-obvious results—the observer calls the story an example of emergence.

In 2006, the theme of **interaction** was explored at the São Paulo Biennale of Art and Technology held at Itaú Cultural under the title “Emoção Art.ficial 3.0 - Interface Cibernética”. In 2008 at “Emoção Art.ficial 4.0 - Emergência!” the theme was **emergence**. Here in 2010 we explore the theme of **autonomy** at “Emoção Art.ficial 5.0 - Autonomia Cibernética”, and so complete the insightful trilogy conceived by Itaulab.

In this text, we first describe the relationships and flow of all three themes in the trilogy. Next, to speak thoroughly of autonomy we introduce a second trilogy of **conversation**, **entailments**, and **autopoiesis**, which provide crucial models along our path. Our journey ends, at least for the moment, by looking at how our two trilogies point to a third: **consciousness**, **meaning**, and **being human**.

THE FIRST TRILOGY: ACHIEVING AUTONOMY FROM INTERACTION AND EMERGENCE

interaction

The organization of a dynamic system circumscribes what it is capable of—what it responds to in the world and what actions it is capable of taking. If we are cyberneticians, we are focused on systems that additionally have a goal and that use their sensing and acting to strive to achieve that goal. The simplest such system is a first-order feedback loop. So: acting, sensing, comparing, acting, sensing, comparing...

Very simple, but amazing. How, we might ask, could such a simple system exhibit an outcome that is surprising to us as observers? Put another way, how does complexity emerge?

emergence

The answer itself is about simplicity and is itself simple: the world is not so simple! It is the interaction between the system and the world that produces the emergent behavior, not the system on its own. The complexity lies in the interaction between the two. So, a simple system surprises us via emergence.

But we may crave something more—so what more could there be? We may improve the system by adding sensors or more internal processing or more ways for a system to take action. But this added complexity may be uninteresting if the system itself, once extant, never changes: our fancy system is merely reactive, responsive, **automatic**—action-in leads to action-out, same every time, without change. How could the system avoid being purely responsive?

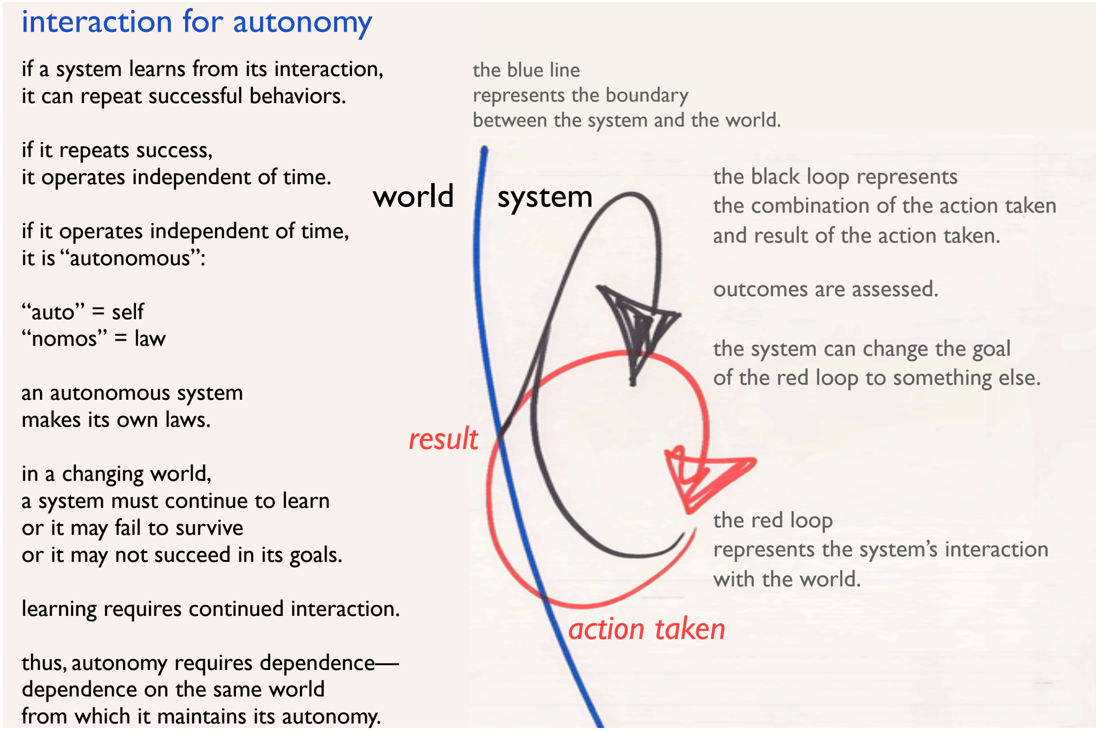
learning

Let us crave that our system **learns**. To learn means to take something from prior circumstances that can be used in future circumstances that are similar, in order to achieve a goal. Such learning may save time, save resources, indeed save the system itself—in the sense that it enables the system to survive in circumstances where, without learning, it would not.

The simplest way to create a learning system is to add a loop of a particular kind: we place the first-order loop inside of a second-order loop that can change the goal of the first one (see sidebar explanation and also Figure 1).

[Sidebar:] Example of a learning system: When a mouse senses a piece of cheese, the mouse's goal should be to eat it; otherwise, without food, it would die. The inner (first-order) loop takes care of getting closer to the cheese (not further away), munching on it when close (not just sniffing it), and other adjustments. But what if the mouse senses a cat? The goal should be to get away; otherwise, it would be eaten. The inner loop takes care of getting further away from the cat (running and dodging, depending on what the cat does), finding a place to hide, and other critical adjustments. So, to know which is which (cheese or cat) and to switch goals accordingly, is a form of learning. This learning may take place by the individual mouse (via trial and error) or by the entire species (via the selecting-out of mice who don't make the right choices). In sum, the only mice around long enough to reproduce are those that are smart enough to know when to eat and when to run. And this can be modeled as an outer (second-order) loop, which gives the system its autonomy.

[Figure 1: Internal organization of an autonomous system that learns.]



As a result of its organization, our double-loop system has created a **memory** of what to do in changing circumstances. (Now or in the future, absent learning anything new or different, the mouse will behave the same way.) Put another way, this memory or learning is *independent of time.* Put another way still, the system has achieved **autonomy** so far as time is concerned.

simple autonomy

"auto" = self. "nomos" = law. An autonomous system makes its own laws.

All learning systems are independent of time, and they may also be more or less independent of a range of possible disturbances in their world. Thus, systems can be said to possess a "degree of autonomy" in relation to different dimensions and different amounts of independence. The more variety in a system's behavior, the more independence it will have from external forces. Thus, a mouse's autonomy will be more than a protozoa and less than a human's.

But we might ask, can a system evolve its autonomy, evolve its laws? If yes, then how does a system's organization change, and even come to increase in complexity over time? How does this evolution exhibit more and more complex **emergent behavior** over time?

THE SECOND TRILOGY: MODELS OF SYSTEM EVOLUTION

Next, we outline a set of models to explain how a system may evolve to increase its degree of autonomy: **conversation** is the process of interaction, **entailments** capture the emergent internal structures, and **autopoiesis** defines the organization of the autonomous system that supports this evolution.

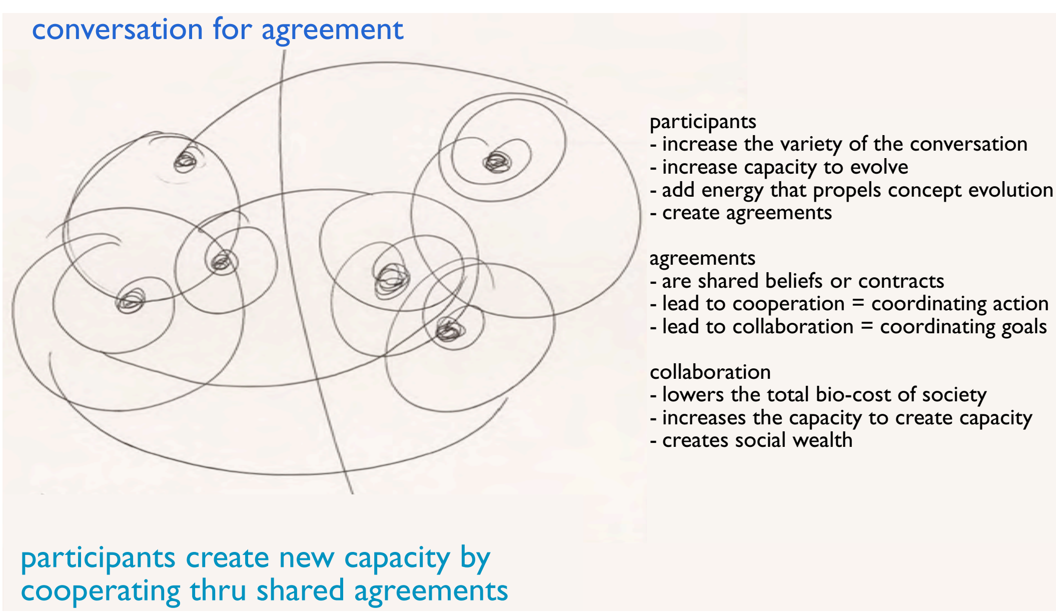
interaction via conversation

As defined by Gordon Pask, conversation is an interaction by which a system receives triggers from the world, interprets them, and changes its internal organization—which is to say, the system creates memories of what worked and didn't, and thereby **learns**. [1] It may learn by changing its internal configuration, which determines how it interprets triggers and subsequently acts. In people, we call this internal configuration a set of "beliefs". A system may evolve its ability to act, as well as its ability to sense the world by growing new sensors [2].

Significantly, some systems that receive a trigger can choose to ignore it, take it lightly and change little, or take it strongly and change more. In other words, a conversational interchange does not force the system to respond; instead, the system takes triggers as contingent and "decides" whether and how to react. (Not all systems have this option; a light switch, for example, cannot say, "I will not be switched!") Systems that can be independent of a trigger, but also able to react to it usefully, possess **conversational autonomy**. The variety of the system defines the range of its possible responses, which may be internal system changes or triggers returned by the system back into the world and to other participants.

Conversation that is useful for all participants requires a common context, a language that is sufficiently shared to begin a useful interaction, and a sweep of exchanges that has some result. The result may be an internal change in one or more participants; shared agreement about an understanding or belief; or even a contract to coordinate actions (for example, to play a game or to exchange money for services). In cases where participants in a conversation are human beings, and because human beings are not predictable, there may emerge highly surprising actions from the conversations. (see Figure 2).

[Figure 2: A conversation may have rich, surprising outcomes.]



But these are the external behaviors. Our goal is to describe the internal evolution, the changes in the system that occur as a result of conversation, especially changes in **belief**. How can we model such emergence, so often surprising and unpredictable?

emergence via entailments

To model the outcomes of conversation such that our view of emergence is clear, we must first return to the simplest form of autonomy, that of independence from time. From there we can elaborate on the evolution of a system to more complex forms.

As stated above, to operate independent of time a system needs a **memory**, a means to retain and hold on to a configuration over time. In the case of organisms with nervous systems, we know there is some means of doing this, even if we don't know precisely how it occurs.

Reflecting on the circular organization of the nervous system, Warren McCulloch wrote in the 1940s, "...activity may be set up in a circuit and continue reverberating around it for an indefinite period of time, so that the realizable [function] may involve reference to past events of an indefinite degree of remoteness." [3] In other words, the system transcends time.

Whenever we "re-member" something, we evoke this reverberation, no matter how remote the interactions that led to this knowing. From that beginning, as we spend more time to understand a concept, the concept itself becomes more consistent and persistent: the experiences converge to become a stable concept. Heinz von Foerster realized that reverberating nervous nets that converged and held memory were faithfully modeled as Eigen Functions. These are mathematical processes that repeat their operation recursively until they converge to a stable point. For example, taking the square root of any number, and then the square root of the result, and then the square root of the result of that, and so on.... the result converges on the value 1. The more it operates from there, the more it simply stays where it is. This stability is memory. [4]

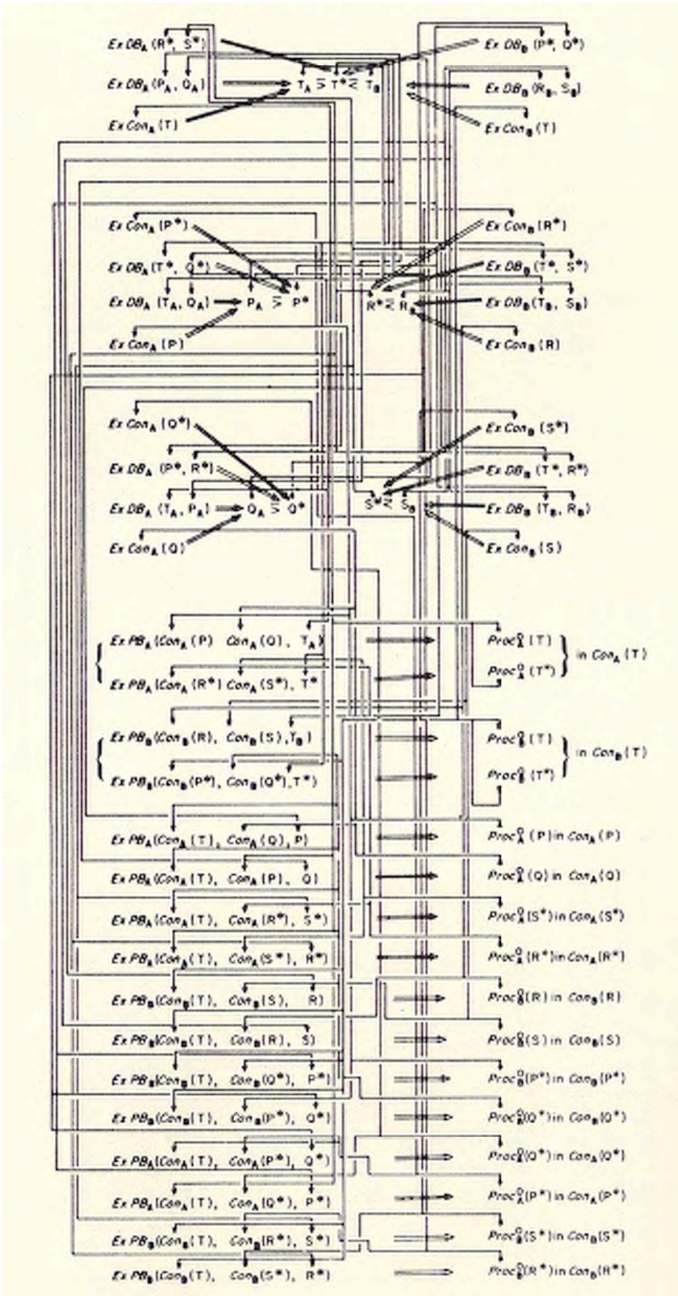
But clearly the memories that emerge from a system are far, far more complex than a single number. How does this complexity emerge? Again, Pask has a proposal, that of entailment meshes, to model the evolution of concepts and beliefs.

Entailment meshes were developed first as representations of understanding. [5] Unlike any other knowledge representation scheme, they require an interdependence among a necessary and sufficient set of distinctions in a mental repertoire such that they are **coherent**. This means that they "make sense together" both descriptively (why) and operationally (how). The canonical example is that of circle, compass, and plane— we may think they have meaning individually, but in practice they only make sense when operating together. (After we understand a distinction and have a name for it, we forget that we learned it only through relationships to other things. It is the very stability or consistency of remembering that fools us into thinking we could know "circle" independent of anything else.)

While this hardly provides a complete explanation of entailment meshes, we have enough to continue our journey by asking the question, how did these configurations of dependent topics themselves arise? And how would new beliefs and further complexity arise from them?

Let us build on the model of memory as reverberating, recursive activity in the nervous system. We conjure a mechanism that starts from existing distinctions and their inter-meshing, combines them with new triggers, and both deconstructs and reconstructs a new internal organization, representing the new belief structure. While this may sound like vague hand-waving, Pask has offered a detailed model [see Figure 3].

[Figure 3: The process synchronization that results from conversation, precursor to the emergence of new distinctions in a mental repertoire.]



My intention is not to explain the figure in any detail, but to point out its intricacies on a path to expressing awe that the process can work at all.

Operationally, Paskian conversation is the synchronization of *a priori* asynchronous processes. This figure shows an intermediate state as two autonomous systems (left side and right side) come to agree on a common understanding in the simplest possible case. [6] Pask proposes that each of the transformations shown by an arrow is necessary for two systems to reach a simple, stable agreement. Before the conversation, each side can only look at the same distinction from a perspective different from the other; after the conversation, both sides can hold either perspective. Thus, their individual belief systems increase in complexity by a substantial amount, by an entirely new perspective.

It may be difficult to imagine how such a complex process is achieved—no less how we might reproduce it in some other embodiment. But it is plausible that trains of electrical impulses have the flexibility, and surely the nervous system has sufficient topology, for doing this. Perhaps the most helpful analog is **resonance.** From the partial harmonies and disharmonies that arise during the conversation, there emerges a more layered but consonant organization that is shared, at least partially, by both sides.

It makes sense intuitively that adding perspectives to an existing organization of beliefs should require so many transformations: the complication of adding new understanding to old is simply daunting. It also shows why it is so difficult for some to change their beliefs. In both cases it is the energies of resonance and dissonance that propel, or prevent, the system's evolution.

So, at one level, we say that conversational interaction affords changes to internal belief. At another level, we say that the nervous system accepts triggers that modulate trains of electrical pulses in the assemblages of concepts in a mental repertoire, resulting in an increase in complexity, and the emergence of new concepts.

Somehow the nervous system is just made to do all this.

autonomy via autopoiesis

Autopoiesis is the term coined by Maturana, Varela and Uribe to describe the nature of living systems as "self-making" [literally "*auto" + "poiesis"*]. [7] It was necessary to coin a new term in order to capture a new meaning, namely, that all the processes involved in the continued existence of a living organism are necessary, nothing is unnecessary, and, all together, everything operates to create a unified presentation of the whole to the world. That is, the system computes the system's (or organism's or person's) boundary.

While we wish to avoid the controversy as to whether autopoiesis applies to social systems, we want nevertheless to apply the concept directly to the organization of cognitive systems. The distinction between "structural" and "interactive" autopoiesis corresponds well to what Pask meant when he described his models as "organizationally closed and informationally open". By the first clause, he meant it is an autopoietic system: it is made up of, and makes itself up from, a repertoire of necessary and sufficient processes (a.k.a., topics), which in turn presents itself as a unity to its peers in an ecosystem (a.k.a., a concept). By the second clause, he meant that it is capable of interactions (conversation) that may lead to change in its organization; that is, it may learn—which has been a motivating strand for our journey here.

THE THIRD TRILOGY: ON BEING HUMAN

With all the foundations laid in our first two trilogies, finally we explore their implications in our third and final trilogy: **consciousness**, **meaning**, and **being human**.

interaction –> conversation –> consciousness

What is consciousness? The word derives from the Latin word meaning "having joint or common knowledge with another". If "to know" means to posses reverberant processes that hold memory and belief, then "to be conscious" means to experience the reinforcement of that resonance with another, through conversation.

Pask conceived his "conversation theory" as a true scientific theory and so required that it have a conservation principal, just as physics has the conservation principals of matter and energy. Neither energy nor matter can be destroyed. For Pask, what is conserved in conversation is *consciousness*, that is, concepts that are shared. [8] Once thought, never forgotten—by the collective whole, at least. Even if some specifics are lost, the *implications* of what has been thought are still found in the form of the remaining systemic organization. It has an effect on what can (and cannot) be thought subsequently.

Surprising as *that* may be, there is a more startling implication that Pask would insist on communicating when he concluded his lectures. Because of the social nature of human beings, because humans share consciousness about concept repertoires—that is, we share what we believe—humans are the medium by which concepts reproduce and evolve. He meant this not at the level of a single human, but rather in a distributed model where human brains are *distributed concept processors*. Stated as an extreme, we are *nothing more* than the means by which concepts exist. We serve as the medium of concepts just as our intestines serve as the medium for *E. coli.*

We thought, by way of our thinking, that we know that we're here (Descartes). Instead, our thoughts do the thinking, and that's why we are here.

emergence –> entailments –> meaning

What is meaning? Why do some things mean more than others? Translating this for an autonomous system: when a trigger arrives, what value does it have for that system?

We need a model for the impact that a trigger may have on a system that is organizationally closed but informationally open, such that meaning can be measured as the "degree of impact" of the trigger. Pask's model of entailments and synchronization as a specific configuration of organizational **resonance** fits perfectly here. Some triggers will reinforce a pre-existing resonance in the circuits of the system, adding strength to the reverberations in the recursive loops. (For example, if we share the response to an experience with someone we care about and they agree with us, then that experience grows in meaning for us.) This increases the persistence of a given memory, a.k.a. belief. Some resonances, such as those described in the section on entailments, cause change in the organization—in other words, a co-evolution of belief. (If we have a deep argument with someone we trust, we may alter how we look at the world forever.)

The greater the reinforcement or the greater the change, the greater the meaning.

Humans are meaning-making systems. Moment to moment we make meaning of our experiences in order to operate in the world, to achieve our goals, to survive. "Intelligence organizes the world by organizing itself." (Jean Piaget). Making meaning helps us to feel more at home in the world: we lower uncertainty, which lowers stress, which makes us more open to new possibilities.

Of course, there is rational meaning and there is emotional meaning. The hormonal system notwithstanding, perhaps the nervous system computes both types of meaning, as implicated in recent findings on "mirror neurons" that fire in the brain under two different conditions: when an animal performs a specific action and when it observes the same action performed by another. Is not this functional example of empathy a form of "emotional resonance"? And wouldn't this resonance be something we notice and reflect on, and create language to express? Here we touch on shared self-knowing, a.k.a. consciousness.

autonomy –> autopoiesis –> being human

What does it mean to be human? Humans live in a social fabric, what Maturana calls structural coupling in language or **languaging**. We thrive—as individuals and as a society—because we live together *in* language.

If Maturana is right to emphasize that *Homo sapiens sapiens* are unique because we live together in language, then we do our living *through* the world, not *in it*. This is because the world is merely the medium through which we converse. I think this gives systems with consciousness a survival advantage, but not everyone agrees.

Either way, with autonomy comes responsibility, for we must choose what we do and decide who we are and what we want to become. Maturana calls this “metadesign”. He insists that technology does not determine us, and that we can choose what we want to be. “Technology is not the solution for human problems because human problems belong to the emotional domain... It is the kind of human being [we want to be that] determines how we use [technology] or what we see in it.” [9]

epilogue on belief

I admit that I do not invest in “collective consciousness” (lacks mechanism) or the “hive mind” theory of the internet (lacks methodology). I am impressed with the brilliance but skeptical of the basis for theories of consciousness in matter. But when I see the detail and coherence of Pask’s theory I am thrilled (the word itself has a resonance and degree of meaning of its own).

Whether as a whole or as individuals—for even individuals are social creatures that cannot stand alone—“we are what we think.”

I can even imagine lashing the “internet of things” to our pre-existing distributed mind, creating a larger incubator of concepts, the true “cyborgian” mixture of animal and machine. But for the moment, the only mechanics that work for evolving entailments are possessed by “meat machines”, as Marvin Minsky, the “father of AI”, has called our brains. Remember, though, that AI architectures never fulfilled their claims to make an intelligence that would organize its own world. But with Pask, we have a detailed treasure map for something more, for a “real” cybernetic space of complex emergence, unifying brains and computing machines, and creating shared responsibilities. To live in conversation, to understand the ethics of our autonomy, to be conscious of our responsibility. To interact, to emerge, to be human together.

Only our ability to resonate limits our participation.

The meaning we make together determines who we are, and who we can come to be. Meaning-making is itself emergent. In the beginning was the interaction, leading to simple emergence and then autonomy, leading to more complex emergence and the autopoiesis of living systems—mental repertoires or physical organisms alike.

Our journey of ontogenesis, of meaning, is thus complete. The story of the trilogy **interaction—emergence—autonomy** from Itaú Cultural’s Emoção Art.ficial 3.0, 4.0 and 5.0 is the co-evolutionary story of all living systems. The insightful trilogy brings us, well, to everything. For being human is all, and everything, that we can be.

The author wishes to acknowledge the autonomous participants in the conversation that led to this approach—Marcos Cuzziol, Guilherme Kujawski, Peter Cariani, Patricia Clough and CJ Maupin—while acknowledging all errors and weaknesses as his own.

references

[1] Cariani, Peter (2009). “Strategies for creating new informational primitives in minds and machines”, Proceedings, Dagstuhl Seminar on Computational Creativity, Schloss Dagstuhl – Lebnitz Center for Informatics, Wadern, Germany.

[2] Pask, Gordon (1975). *Conversation, Cognition and Learning*. Amsterdam, New York: Elsevier.

[3] McCulloch, Warren (1943). “A Logical Calculus of Ideas Immanent in Nervous Activity”, reprinted in *Embodiments of Mind* (1965), Cambridge, MA: MIT Press, pp. 40-45.

[4] von Foerster Heinz (1973). “Objects: Tokens for (Eileen-)Behaviors”, reprinted in *Understanding Understanding*. New York: Springer, pp. 261-271.

[5] Pask, Gordon (1976). "Conversational Techniques in the Study and Practice of Education", *British Journal of Educational Psychology*, Volume 46, I, pp. 12-25.

[6] Pask, Gordon (1980), "Developments in Conversation Theory", *International Journal of Man-Machine Studies*, Volume 13, pp. 357-411.

[7] Varela, Francisco J.; Maturana, Humberto R.; & Uribe, R. (1974). “Autopoiesis: the organization of living systems, its characterization and a model”, *Biosystems* 5 pp. 187–196.

[8] Pask, G. (1981). “Organisational closure of potentially conscious systems”, in *Autopoiesis*, M. Zelany (Ed.), New York: Elsevier, North Holland, pp. 265-307.

[9] Maturana, Umberto (1997). “Metadesign”, available at http://www.inteco.cl/articulos/006/texto_ing.htm.

-end-