Human-Centred and Creative AI in Performative Robots

Authors:
Prof Craig Vear cvear@dmu.ac.uk
Dr Fabrizio Augusto Poltronieri fabrizio.poltronieri@dmu.ac.uk
Institute of Creative Technologies
De Montfort University, Leicester, UK

Introduction

This chapter discusses two performative robot projects that share a core ethos about the nature of human-centred and creative AI: *Embodied Robots for Music* (2018 – present) by Craig Vear and *Amigóide* (2010 – present) by Fabrizio Augusto Poltronieri. Human-centred AI (HC-AI) focuses on the design, development and deployment of intelligent systems that co-operate with humans in real-time in a 'deep and meaningful way' (MIT 2019). HC-AI is 'defined by two goals: (1) the AI system must continually improve by learning from humans while (2) creating an effective and fulfilling human-robot interaction experience' (Ibid.). The two projects discussed in this chapter both apply these core goals as a central philosophy from which the concepts of Creative AI and Experiential Learning in the context of performative robots are developed. At the centre of this discussion is the articulation of a shift in thinking of what constitutes Creative AI and new HC-AI forms of machine learning from inside the flow of the shared experience between robot and human. The first case study is *Embodied Robots for Music* which investigates the technical solutions and artistic potential of AI driven robots co-creating with an improvising human musician in real-time. The second *Amigóide* performs a range of functions according to a predetermined set of coded instructions in search of humans to engage into friendships. These projects are ongoing and are part of the Creative AI Research Group in the Institute of Creative Technologies at De Montfort University[1].

Case Study 1 – Embodied Robots for Music

The goal

The aim of this practice-based research project is to investigate the technical solutions and artistic potential of AI driven robots co-creating with a human musician in realtime. This is an ongoing project that was initiated in 2014 with a conference *Embodied Intelligence in Music*[2] that investigated the meeting point of embodied cognition, artificial intelligence, music composition and performance, game and software philosophy, in order to develop the parameters of a new field of research.

The overarching research question with *Embodied Robots for Music* is: *If we want robots to join us inside the creative acts of music then how do we design and develop robot systems that prioritise the relationships that bind musicians inside the flow of music-making?* This question comes from the deep and meaningful experiences I had as a professional high-level musician for over 30 years and supports one of the core goals of HC-AI. Overall, I am curious about the technical innovation to make this happen; but more so by the new avenues of creativity that can open up to us humans. The overarching goal of this research is not to make machines more creative, but to make humans more creative through deep and meaningful relationships with co-operating intelligent machines.

Definitions

Before I describe the solutions that I designed and deployed in the *Embodied Robots for Music* project I need to define in simple terms what I mean by the following terms within the context of this project. This also

[1] https://www.ioct.co.uk/creative-ai
[2] https://eim-live.our.dmu.ac.uk

helped consolidate that the design and development of the Creative AI and the robotic systems with common goals.

Musicking – the creative acts of realtime music-making. Musicking is a term first created by Christopher Small to define a perspective that 'to music is to take part' (Small 1998). Small wrote that taking part can happen 'in any capacity' (ibid.) such as performing, composing and listening. Crucially meaning formation through musicking is formed in the relationships that are established within the realm of taking part with spaces, agents, sounds, and presences that are encountered here.

Flow – is the experience of musicking from inside the activity. Within the context of this project, the flow of musicking defines how 'musicians become absorbed in the music through a sense of incorporation within their environment (the soundworld), a shared effort (with the digital, virtual, AI and robotic agents) and a loss of awareness of their day-to-day wakefulness and bodily self-consciousness (embodiment with their instrument and into their music) (Vear 2019: 192).

Embodiment (in music) – the process in musicking of drawing the musician's sound into their bodily sense-of-being. This presumes that when musicians make music it is not a process of outputting sound into the world, but an embodied experience of *becoming* the sound they create in the flow of musicking. Equally, it describes the process of the musician reaching out from this sense of becoming and drawing in the sounds of others so they feel their *presence-as-sound*. This is a dance of sorts: to touch, to feel, to sense, to work with, to play with, to hide and seek and flirt and subvert, with others through the *flow*.

Creativity & Co-operation - when play turns into invention within the flow of musicking. However, creativity through play-involvement is not a constant, nor is it reliable or automatic; it needs nurturing with open, generous and cultivating energy. On the other hand, it can sustain bold and mischievous challenges or seemingly disruptive engagement designed to rail-road on-going trains-of-thought - so long as these are still giving in their nature. Within this project I define three sub-domains of creativity to highlight the human-robot relationships. These are based on my general experience as an improvising musician, and are used to identify the types of co-creativity within musicking from the human musician's perception (note – this project does not deal with notions of machine consciousness or perception):

- Concurrent – a sense that both agents (human & robot) are playfully inventing in isolation but within the shared flow of musicking
- Collaborative – a sense that both agents are contributing to a shared play idea, feeding a sense of collective invention through individual contribution and perspective
- Co-creative – a sense that the robot and human agents are collectively inventing through a stimulated sense that each are in inside each other's head. By that I mean that the relationship, as perceived by the human musician, is that the robot/AI is in-the-loop with the human and together they are inventing on a singular idea, feeding each other's play as if it were one train-of-thought.

Creative AI - includes practices that are self-defined as such, and therefore have AI embedded into the process of creation, but also encompasses novel AI approaches in the realisation and experience of such work.

With these definitions in mind the goal of the *Embodied Robots for Music* project is to design and develop a Creative AI system that enhances human musician creativity by stimulating, inspiring, interacting and co-operating in the flow of embodied live improvised music-making. Therefore, to build a robot driven by a Creative AI system, it must (1) continually improve by learning from humans while (2) create an effective and fulfilling human-robot interaction experience (MIT 2019).

[Not] The Solution

Before I describe my solution, I'd like to describe what it isn't. First, it is not an instrument performing robot. For example, *TeoTronico* (2012) is a pianist-robot, designed and built by Matteo Suzzi. This robot plays the piano with dynamic control and articulation, moving 53 levers (described as fingers by Suzzi) with 'great accuracy and speed' (Prosseda 2014). In one example on YouTube it plays a piece composed by Mozart, extremely well. It seems to have a sensitivity about its performance and, even though the designers state that it uses MIDI files or a 'mirror pianist' (ibid.), it does not sound like it is driven by a standard quantized MIDI file so some form of human capture was used that stored a human performance as a MIDI file, which *TeoTronico* replayed. Its flow has been pre-packaged, and then regurgitated. Its sound is in the now, but its musicking is not responding to the now, nor its environment. As such, to achieve the main aim of the *Embodied Robots for Music* project a technical and artistic approach such as this pianist-robot would fail as it simply could not co-operate with the human.

Secondly, it is not a goal-specific humanoid-robot. For example, environment aware and goal-cognizant robots such as those being developed for the human-robot World Cup in 2050[3] are sophisticated robots employing AI that make them aware of their world. In general, these systems use computer vision and sensors to navigate through this world; they have real-time awareness of the here-and-now and are interacting with that in their goal to get the ball and score. The problem of using this kind of approach with the *Embodied Robots for Music* project is in the nature of embodied interaction. In musicking, the embodied relationships with other musicians are with the presence of the others as sound, not with them as human-flesh. This relationship goes beyond relating to their physical presence and their movement, although it does play a part in varying degrees, and at varying times, in my on-going relationship-building process. So, an *Embodied Robot for Music* needs to create relationships with human musicians through its *presence-as-sound*, yet also have some physical presence and movement to inform this. Using these Football-robots as an analogy, it is not the physical movement of the robot moving towards the ball, or kicking, that creates the relationships required for this project, but the relationship with the flow of the movement of the ball. As such, it is not the movement that incites the sound that is being related to in musicking, but the presence of that sound in the flow.

Lastly, it is not a deep learning AI robot. Using deep learning and neural networks to teach a robot to operate like a musician does not yet feel like a productive way forward. The problem is not the process of machine learning (as is described below the embodied musicking robots and AI system utilizes machine learning and does use a dataset), but the process of learning what. The current state-of-art is focused on training networks and AI to achieve knowledge and reasoning of a certain task or application. For example, in 2016 Google's DeepMind researchers developed *AlphaGo* which was the first computer Go program to beat a human professional Go world champion without handicaps on a full-sized board. *AlphaGo* uses complex tree search algorithms to determine its next moves based on 'knowledge previously "learned" by machine learning, specifically by an artificial neural network (a deep learning method) by extensive training, both from human and computer play' (Wikipedia 2019). The focus of learning in these systems – and many others like it – in on producing some thinking-mind that can operate logic and reasoning of a single task/application at that comparable to, or better than a human. But embodied interaction in musicking and a human-centered approach does not rely on notions of human intelligence based on logic and reasoning, therefore seeking a solution for an embodied robot that co-operates inside musicking, using human-level intelligence and reasoning is a blind alley: it's actually too complex, and too little is really understood about it, and it's too Cartesian.

The solution

The solution to the research question posed above, needed to avoid the above issues and to design, develop and deploy a robotic Creative AI that would have a presence within the co-creativity of the flow of musicking, and not be an AI-zombie. This approached reinforces the personal understanding that when a musician enters the world of musicking, the 'I' is coping in a very different world-of-concern than if they

[3] https://www.vice.com/en_us/article/nzep7q/robocup2015

were walking down a street. In a sense 'I' becomes a different creature with a different set of priorities and concerns, outlooks and sensorial inputs than my normal, human wakefulness. The technical and artistic solution for *Embodied Robots for Music* focused on a robot that was first and foremost a coping entity in this specific world-of-concern (the flow of musicking).

The solution was to develop a system based on these three principles:

1. Coping – an *Embodied Robot for Music* needed to cope in real time within the realm of musicking, and be present as sound whose movements are embodied within such flow. This required a non-representational approach to how it related to the flow as the coping mechanisms needed to be open and dynamic enough to co-operate in any given musicking realm. Limiting the robot to a single representation of what musicking is, or might be, imposed onto the system by the human designer(s), would only work in a number of instances.
2. Creative AI dataset and Experiential Learning – these concepts needed to be designed from within the realm of musicking, prioritising the phenomena of being inside this realm, and capturing an essence of what it means to be embodied within the flow. Current techniques in deep learning would not support this priority, so a new concept of Experiential Learning was needed to be designed (discussed below)
3. Belief – the robot needs to believe in its view of the musicking world through limitations, embedded aesthetics and behavioural traits (even glitches and bugs in the system)

From the human-centred artistic perspective this *Embodied Robot* needed to address the following

- The robot was not an extension of the musician; but should extend their creativity;
- The robot should not be an obedient dog or responsive insect jumping at my commands or impetus, but a playful other;
- It should not operate as a simulation of play, but as a stimulation of the human's creativity;
- It is not a tool to enhance the human's creativity, but a being with presence in the world that they believe to be co-creating with them, and
- It should prioritise emergence, surprise, mischief, not expectation.

Technical Solution

Coping

The design of the *Embodied Robot for Music* was informed by two early papers by the robot innovator Rodney Brooks, specifically *Intelligence without Reason* (1991) and *Intelligence without Representation* (1987). In these he lays out the foundation of his approach to designing and building robots that are first and foremost able to cope and therefore adapt to a dynamically changing environment within the parameters of specific and multiple goals. This research eventually led to his robots being used for space and sea exploration, military and medical application and the iRobot Roomba vacuum cleaner series. These Roombas are designed with *iAdapt* AI to be 'creatures' that cope in a specific world-of-concern; in real-time. They do not have a model of representation of their world (such as building a 3D model of the space through computer vision and object analysis), nor do they make one as it goes about its business, but uses goals and strategies to cope with whatever that world can throw at it (furniture, steps, chairs that get moved one day to the next).

Brooks' foundational theories, and observations of my own Roomba, guided the developed for my *Embodied Robots for Music*, and generated this set of principles (adapted from Brooks 1987):

- An *Embodied Robot for Music* must cope in an appropriate musical manner, and in a timely fashion, with the dynamic shifts inside the musicking world.
- An *Embodied Robot for Music* should be robust to the dynamic environment of musicking, it should not fail to minor changes in the properties of the flow of musicking, and should behave appropriately to its ongoing perception of the flow.
- An *Embodied Robot for Music* should maintain multiple goals, changing as required and adapting to its world by capitalising on creative opportunity.

- An *Embodied Robot for Music* should do something in the world of musicking, ‘it should have some purpose in being’ (Brooks 1987).

The multiple goals were (in order of priority):

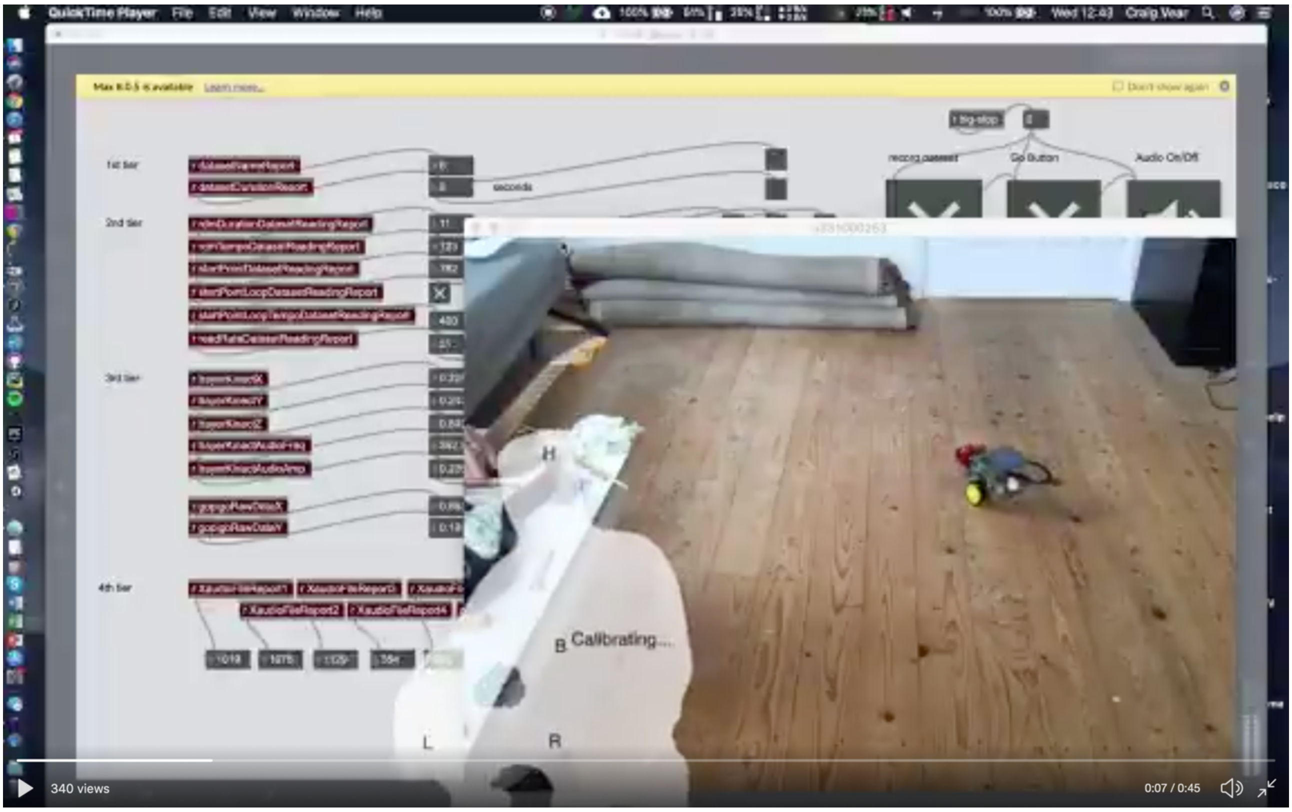
I. Self-preservation – the robot must avoid obstacles, not crash into the other musician or fall off the stage.

II. Instinctual behaviour – if left alone the robot would make music. This was driven by the Creative AI dataset (discussed below), which operated as its DNA of musicking creativity.

III. Dynamic Interaction – the robot can, in certain conditions, be affected by the sound of the live musician. Using a process of simulated affect linking (Gelertner 1994) the Creative AI could leap between related, abstracted or unexpected datasets. Metaphorically, the robot’s internal trains-of-thought would be triggered by phrasing (short-term temporal limits) and the dynamic impetus of the human.

IV. Short-term memory – the robot would have a memory buffer of the live music, which would influence its responses and operate as a thematic repository to be called upon as an option.

A critical feature of the design of *Embodied Robots for Music* is that each of these goals directly move the wheels. It was essential that each goal is not part of an elaborate, logically-flowing, representation of a thought process, mimicking some-kind-of-mind. As such, the overall design of robotic system was modular, with each system directly accessing the wheels when in operation.



Creative AI dataset and Experiential Learning

This project innovated a different approach to machine learning that involved a human in-the-loop and an in-the-groove approach. In contract, to the current techniques in machine learning, this Experiential Learning (EL) approach trained the AI on-the-job and crucially inside the flow of embodied musicking. Furthermore, that the EL process collaborated with a human musician who was equally learning about this new musicking system. This approach supported both the human and the AI to automatically learn and improve from experience.

The EL process focussed on capturing the physical phenomena of an improvising human musician in the flow of creative musicking. The sensing mechanism used 3D depth tracking of the human musician's body using a Kinect sensor (simply x, y, z movement of both hands, body centre and head), and Fast Fourier Transfer (FFT) analysis of the live sound (fundamental frequency and amplitude). The resulting dataset reflected the position and rotation of an embodied musicking body in motion together with the amplitude and frequency analysis of the actual sound made by such movement, without preserving the performer's mass, musculature, melodic shape or music. Thus, embodied musicking movement is extracted from the performer's body while they are making music; in a poetic sense the dataset contains the meta-level DNA of musicking without the specifics, or a representation of the music. Recorded audio-video capture of the music performance would always anchor the dataset to a specific person and point in time, whereas the meta-level data could become the building blocks for the virtual composition. As data phrases can be edited, treated and re-purposed by the robot's AI again and again without the risk of repetition.

The initial process involved seeding the Creative AI dataset by capturing the live performance of an improvising musician. Once a small set had been generated, this musician then worked with the robot through a series of training sessions, with this new live data being added to the Creative AI dataset. The more they worked together the more meta-level DNA of their shared creativity would be put back into the dataset system, thereby improving the AI's knowledge base of experience of embodied musicking.

The EL approach learns through embodied interaction inside the flow of musicking. It prioritises the capture of the meta-level DNA of its improvising partner – the human musician – and extracted elements from the dataset (any data randomly chosen by the system as it is all endowed with meta-level creativity) into its AI processing, and then outputs the resultant sound as music. This EL process enhances the dataset through experience by its embodied coping inside the flow of musicking. The human musician perceives meaning in the robot's musicking who in turn co-operates in the making of music (generally perceiving the relationship through one of the perspectives of creative co-operation discussed above) and responds with a creative solution through music. This is then captured using the sensing mechanisms and stored back into the Creative AI dataset. Thus, the cycle of EL continues to enhance and improve the dataset and enlarging the Creative AI memory bank of deep and meaningful interactions between human and robot which in turn forms the basis for future interactions.

Belief

It might seem odd to implement belief into the AI of a robot, given that this term usually refers to a religious or spiritual faith, but, it is the broader definition of this word that I am particularly interested, specifically that *Embodied Robots for Music* have an acceptance that something is true, or that it has trust or confidence in something. In short, the robot's operational systems are embedded with structures that it can accept as guiding beliefs. These are:

A) *Movement behaviour* – The robot's movement operates within a behavioral system, designed to react openly to the dynamic soundworld and move the wheels accordingly. The robot AI makes choices determined by whichever goal (listed above) is driving the wheels at any given point but within fixed parameters. The *Embodied Robot for Music* has freedom of choice to operate within such a field of response-possibility. These are based on human preferences and outline a range of creative choices which have been determined over a number of decades. These are personal and subjective, and if these parameters where to be shifted of changed then a different set of musicking characteristics would emerge. Within this structure the robot has been embedded with a sense of aesthetic that it can trust (believe to be true) and that the choices it makes are appropriate to it co-creating inside the flow of musicking, and unique to itself.

B) *Soundworld* – the robot has a fixed sound library of roughly 2000 short sounds, that were recorded through a live improvisation, thereby embedding them with an essence of musicianship. These are triggered only when the wheels move. These are then either presented to the world in their raw state or are treated in some way (time stretch, pitch shift, or both) using the Creative AI dataset as controlling parameters. The robot does not have the whole possible world of sounds, synthesis and composition at its finger-tips, it's sounds have a character and an aesthetic basis which it can use to express its behavior and be unique to itself.

C) *Creative AI* – At the core of the Creative AI dataset is a world of embodied musicking captured through the EL process (described above) and through live interaction. This data is used to control every aspect of the AI, movement, sound production choices, and interaction goal. The dataset is also used to make choices about how the dataset is to be recalled and read by the algorithms (e.g. the read rate and ramp speed for each instance of wheel movement). This means that the direct application of data into wheel movement, and also the translations of that into sound-object choice and therefore as music in the flow, is imbued with the essence of embodied musicking that has been embedded in the core of the dataset.

But really these embedded belief structures are there so that the human musician can believe that the robot's behavior and responses are truly emanating through musicking, and to draw attention to that fact this robot is a valuable co-creative presence inside a shared flow. We all know that this robot is really an assembly of plastic and metal components together with a couple of motors and a processor. But because the human musician can trust it believes in certain things and has been embedded with a certain notion of its world-of-concern through concepts such as affectual response, its range of sonic choices, and its behavior, the human musician can believe in it as a co-creative collaborator inside musicking; and that is what happens, and that is what leads to deep and meaningful human-centered interactions..

Conclusion

Using the principles outlined above the *Embodied Robots for Music* project has created a co-creative system that responds to the interaction with a human musician through a cyclical relational process. It is important to note that the interaction with the musician begets movement as its primary goal for musicking, and that this movement is embedded with essence of embodied musicking because of the Experiential Learning process. Following this, the movement begets sound, which begets music such that all relationships between human and AI are informed by phenomenon data captured within the embodied flow of music-making: either from the Creative AI dataset or through live interaction. The ultimate goal of this research is *not* to find solutions to replace human creativity, but to enhance it and move it forward into new discoveries. In short, this research is seeking to find experiences like those emergent through DeepMind and Alpha Go's interaction with the professional Go players. In the 2019 film of this project several of these professionals reflected that when they played with AlphaGo they 'see the world different, before everything begin […] With this thing, I will change something with my game. Maybe he just can show humans something we never discovered. Maybe its beautiful', and 'like move 37, something beautiful occurred there"; "in a broad sense move 37 begat move 78 begat a new attitude, a new way of seeing the game ... he improved through this machine, his humanness was expanded after playing this inanimate creation' (AlphaGo 2017).

Case Study 2 – Amigóide: Engagement, Interactivity and Robotics

Amigóide[4] (figure 1) is an automaton – an apparatus which performs a range of functions according to a predetermined set of coded instructions – which searches for humans to engage in friendship with. It is an award-winning project initially commissioned in 2010 by the Itaú Rumos Prize, one of the most important art prizes in Latin America. Its debut was in a group exhibition in 2011, whose topic was art and cybernetics, held at the Instituto Itaú Cultural, in the heart of São Paulo, Brazil.

Two versions (1.0 and 2.0) have been developed, using different AI approaches. The first one (2010-2011) uses GOFAI (Good Old Fashioned Artificial Intelligence) (Haugeland 1993), whereas version 2.0

[4] The made up word "Amigóide" plays with the Portuguese word "Amigo", which literally means "friend".

(2019) was built utilizing a mix of GOFAI and deep learning, taking advantage of modern machine learning frameworks, which allowed the use of computer vision and image recognition techniques in real-time. This case study describes the process behind the automation's design and developments, including the benefits encountered in the transition from GOFAI to a mixed AI approach.

The automaton interacts with people through movements, LEDs, a pre-recorded synthetized voice and demonstrations of digital feelings, which in this case are very simple, with Amigóide reacting through its lights when the infrared sensor readings (version 1.0) or computer vision analysis of the camera video stream (version 2.0) indicate whether a human interactor gets close to it or not after a round of interaction.

Once the automaton finds a human it starts to follow this potential friend, intending to start an endless, flawless friendship. It gets closer to the eminent friend and poses the question: "Fabian, do you want to be my friend?". Fabian is an imaginary friend, a super one, programmed into the automaton's mind. After the initial contact, Amigóide tries to conquer the friend through a series of phrases such as "Fabian, what do you search in a friendship?"

Amigóide can be described as a rational agent, i.e., an agent "that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome" (Russell and Norvig 2016:4). Hence, Amigóide is a rational agent whose best expected outcome is to establish contact with humans and conquer their friendship. The concept of rational agent (Russell and Norvig 2016) is pivotal to this project because it leaves room for uncertainty. The aim of a rational agent is not to achieve perfect rationality – always doing the right thing, that in this case would be to do everything to acquire lifelong friendships –, as it is impossible in complicated environments (Russell and Norvig 2016).

Figure 1: Amigóide 1.0

Any sensory system cannot provide perfectly reliable information about the environment, and the cheap infrared sensor in version 1.0 and version 2.0's camera are not exceptions to this rule. Uncertainty needs to be handled by some sort of reasoning and planning, embedded in algorithms. In both versions of Amigóide, all reasoning and planning are handled by classic, GOFAI logic. Although it appears simple, Amigoide's environment is sophisticated and highly complex, as it involves processing human actions in real time and in unpredictable ways. Thus, what Amigóide does is to try to act appropriately even when there is not enough time to carry out all the computations and reasoning required to achieve the best possible outcome, a concept known as "limited rationality" (Russell and Norvig 2016).

The automation especially fascinates children and teenagers and the public's participation is crucial. Usually, at the start, the audience doesn't know what to expect. However, once people realize the way Amigóide acts they try to draw its attention. As the automaton tries to follow only one person at a time, the others either keep trying to grab its attention or start wondering how it works and what kind of technology is behind it, displaying different kinds of behaviours, like running after the automaton or screaming answers to its questions, even if Amigóide is completely deaf, unable to process the words said by the human interactors.

After almost every session the audience keeps on discussing and trying to find answers to explain what they experienced, making the project an exciting way to expose people to hard subjects such as artificial intelligence, spatial consciousness, cybernetics, rationality, and cognition, as all of them play an important role in the way Amigóide works.

Physically, the automaton was designed to challenge the idea of human-likeness in robotics, featuring a custom designed plastic cylinder – one of the most basic geometric shapes –, housing 40 RGB LEDs and a loudspeaker as a lid. This elegant solution fits into the plastic base which supports the wheels and the electronics (figure 2).

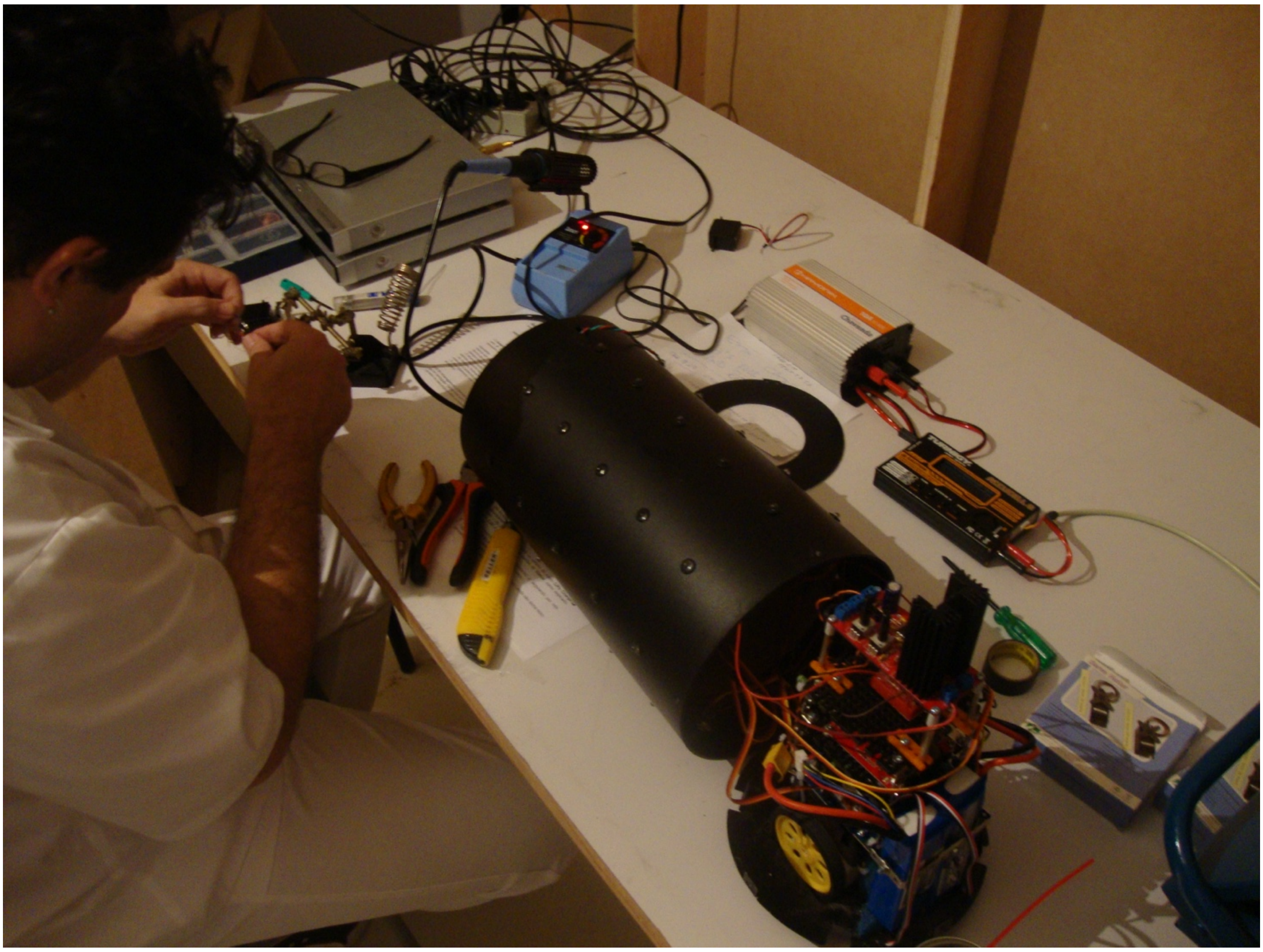


Figure 2: Amigóide 1.0 disassembled, showing its internal parts and the plastic cylinder which serves as its body.

It is worth mentioning that this is a controversial solution for a robot acting as an art installation and designed to interact with people, as an ongoing question in the robotics community concerns the degree of human-likeness robots ought to have when interacting with humans (Riek et al. 2009). This question is usually framed within the context of the "Uncanny Valley", a theory conceived more than 40 years ago by Mori (2012), a robotics professor at the Tokyo Institute of Technology. The author argues that he has "noticed that, in climbing toward the goal of making robots appear like a human, our affinity for them increases until we come to a valley, which I call the *uncanny valley*" (Mori et al. 2012:98). Mori's theory also proposes that as robots become more human-like they become more familiar, and thus more likeable, until the unavoidable "mismatch between their form, interactivity, and motion quality elicits a sense of unease" (Riek et al. 2009:1). With Amigóide, I wanted, since the beginning of its conception, to avoid at all costs any human-likeness, proposing a break with the theory of the mimesis of Plato, where the Greek philosopher explicitly says that all art is mimetic by nature, and art is an imitation of life (Plato, 2007). Amigóide should look as abstract as possible, even if its main goal is to interact with humans, acquiring human friendship in its own way. At the end, my design proved right, as not one person who interacted with Amigóide showed any kind of rejection towards the automaton.

Amigóide 1.0 – A GOFAI approach to interacting robots and their limitations

For version 1.0 – a special room was designed (figure 3), where up to 10 people could interact with it at the same time. The engagement with humans was made via a random selection process, based on the analysis of data from the environment provided by an infrared sensor on the top of the automaton. At the core of Amigóide 1.0 was an Arduino NG that controlled a DC motor, a set of LEDs, an audio interface shield connected to an amplifier and two sensors: the aforementioned infrared one and a colour sensor on the bottom of the plastic body, facing the floor. This colour sensor was used to identify the black border that surrounded the automaton's circular room, delimitating the robot's range. When the colour sensor identified the black strip on the floor, it triggered a routine which made Amigóide rotate and re-enter the room. This black border was necessary due to the fact that no true computer vision systems were used in this version, and it was impossible to distinguish between humans and walls with the simple data supplied by the infrared sensor. This solution prevented version 1.0 from bumping into walls, at the expense of having to consider as a human everything detectable inside the main circle. The system was powered by a heavy lithium battery, designed to power aero-models, which lasted about 6 hours.

Figure 3: Amigóide 1.0 in its room, especially designed to support interactions with up to 10 people at the same time.

Restricted to its partially observable environment, since its inaccurate sensor was unable to cover all the perimeter, version 1.0 had to follow a strict set of rules to have some sort of autonomy and learn something from what it perceived. The environment was designed to be as deterministic as possible, even if it was not a fully observable environment, to try and avoid any kind of stochastic process, meaning that all future actions performed by the automaton were determined by the environment's current state. The logic behind the actions was planned to be episodic rather than sequential. According to Russell and Norvig, "in an episodic task environment, the agent's experience is divided into atomic episodes. In each episode, the agent receives a percept and then performs a single action. Crucially, the next episode does not depend on the actions taken in previous episodes" (2016:44), whereas in sequential environments, "the current decision could affect all future decisions" (Russell and Norvig 2016:44).

Although it was a huge success, Amigóide's first version had several limitations related to the technological choices employed in its construction. We need to bear in mind that this initial version was designed from scratch in 2010, before the actual wave of machine and deep learning algorithms was widely disseminated and available. These technological constrains had a negative impact on the way version 1.0 behaved and the level of interaction with humans it was capable of. It was a simple reflex agent, acting on the basis of its current percept, which informed the series of condition-action rules that handle its behaviour. The major limitations of this approach, however, started to show when I tried to improve many of Amigoide's basic capabilities – such as its environmental sensing – and to implement new features, like the ability to perform in any environment, and not only in a special room.

These problems were hard to be defined by a list of formal hard-coded rules. The fact that Arduinos use C language did not help either, generating a hard to follow and to debug code, with endless if-then-else structures and calls to very low level routines. In the end, these questions proved to be not only technical but rather philosophical ones, intimately related to the approach used to design the automaton. My attempt to use high-level symbolic, human-readable representations of problems was not flexible enough to provide the level of versatility I needed, at least not at a reasonable cost benefit level. GOFAI was holding me back and it took us a long time to realize that. For this reason, and because the history of GOFAI proves that it has been providing a useful experimental framework for the creative use of computers, it is important to have an in-depth discussion about this topic. I also believe that every technology brings its own ideology – the telegraph is a good example (Carey 1983) –, and it is crucial to debate how technological choices impact creative and artistic decisions.

According to Haugeland (1993:112), "GOFAI, as a branch of cognitive science, rests on a particular theory of intelligence and thought – essentially Hobbes' idea that ratiocination is computation". For Hobbes, the reasoning was like numerical computation, in the sense that we add and subtract silently in our heads. GOFAI systems attempt to implement theoretical basis that emulate human behaviour through the construction of software and hardware, such as Amigoide 1.0. Historically, it aims to provide cognitive results at least comparable, and preferably superior, to the expected behaviour of intelligent human beings in similar circumstances, and in all the human cognitive domains, including the abilities to write, read, draw, talk and interact. According to Floridi (1999), GOFAI tries to achieve its goal by seeking to find a balance between two distinct approaches:

1. A Cartesian one, based on a rationalist dualism, where the presence of intelligence is completely detached and independent of the presence of a body, constituting a complete disembodied cognitive system. It would, at least in principle, be implementable by other disembodied and stand-alone kinds of cognitive systems. However, intelligence is also completely mind-dependent – mind here meaning the human mind –, making it impossible for a machine or even an animal to achieve sophisticated intelligence in this manner.
2. A materialist monistic attitude, which considers that intelligence is solely a complex property of a physical body. In this vision, mental processes play a secondary role, being caused by physical brain processes. This view draws support from the philosophical theories of epiphenomenalism and mechanistic materialism. According to this view, mental events are caused by physical events in the brain but have no effects at all upon any physical events, i.e., mental states or events are by-products of states or events in the brain, necessarily caused by them but exercising no causality themselves. Thus, a thought, belief, desire, intention, or sensation is produced by a specific brain state or event but in no way affecting the brain or the body to which the brain is connected. Hence, this approach would, theoretically, be implementable by other kinds of embodied cognitive systems, including computers and robots.

Facing such paradoxes, where the initial propositions in both philosophical views are favourable, but the outcomes are not achievable, GOFAI implements a form of computational materialism, described in the first half of the materialist monistic approach (item 2 above). Therefore, GOFAI should be possible to be achieved by any brainless, mindless and lifeless general-purpose, symbolic system, deprived of psychological or embodied experience, which is not the case, as out automaton demonstrates.
In order to implement consistent computational materialism, a GOFAI based system has to adopt an extreme reductionist and abstract version of computational materialism, assuming the following presupposition:

INTELLIGENCE → RATIOCINATION → SYMBOLIC PROCESSING → COMPUTATION

This reduction led to a type of functional behaviourism, which has many negative implications, including the fact that GOFAI is not scalable – as, again, proved by Amigoide 1.0 – or broadly applicable, even to problems that are in some way similar to the original ones for which the system was designed. The problem I faced with the need of a special room for the automaton is a good example. According to Dreyfus, it has turnout out, after decades of investment in GOFAI, that the "research program based on the assumption that human beings produce intelligence using facts and rules has reached a dead end, and there is no reason to think it could ever succeed" (1992:11).

Although GOFAI has been suffering from recurrent criticisms, I am not claiming that it should be abandoned or saying that it is no longer relevant. I have been using GOFAI techniques in combination with neural networks in my experiments and artworks, as the next section demonstrates. This mixed approach is supported by several AI front-runners, as highlighted by Ford (2018).

Amigóide 2.0 – Deep learning, computer vision and beyond.

After years with the project being shelved, and having worked with deep learning in different artistic contexts in the last few years, I decided to give Amigóide another shot after getting a $99 NVIDIA Jetson Nano board[5], which could replace not only the Arduino used in the original design, but also the MP3 module that was used to play the pre-recorded voices from an SD card in the first version. Running Ubuntu, a Linux flavour, the Jetson Nano is capable of producing a real-time synthesized voice.

One of the reasons I was skeptical about using deep learning and computer vision in a project like Amigóide was the fact that deep learning energy consumption is usually very high, but it is not the case of this board. Besides that, I have many concerns regarding deep learning, and believe that in the near future new technologies and techniques will supplement it. I share many of the concerns about deep learning described by Marcus (2018), who provides a comprehensive overview of the problems around modern neural networks.

Despite my criticism about machine learning and deep learning, it is undeniable that many problems found in GOFAI, especially the ones concerning computer vision and speech recognition, are much easier to solve using neural networks, mainly because they bring a different mindset, a distinct philosophical approach, that is beneficial to the field of Creative AI, notably when mixed with GOFAI techniques.

In contrast to GOFAI, machine learning – of which deep learning is a subfield – seeks to allow computers to "learn" from experience and come to some understanding about the world in terms of a hierarchy of concepts, where each concept is defined in terms of its relation to other simple concepts (Goodfellow et al., 2016). The term "learning" can be misleading, as, at the actual stage of development, machines do not really learn. Machine learn algorithms typically apply mathematical formulas to a collection of inputs – the training data in supervised learning, the most used type of machine learning nowadays – in order to produce the desired outputs. A machine learning model is the result of "brute force" curve-fitting. When the training is successful, the model can map new inputs to the patterns found in the training data. That is the reason why it is expected that these mathematical formulas also generate the correct outputs, or predictions, for most other inputs distinct from the training data, respecting the condition that those inputs are from the same or a very similar statistical distribution as the ones found in training data. It is not proper leaning, in the human sense, because even a slight distortion or change in the inputs is likely to produce completely wrong outputs, what can be interesting for creative and artistic purposes.

The first learning algorithms were intended to be computational models of biological learning, trying to mimic the way learning possibly happens in the brain. According to Goodfellow et al. (2016), deep learning is motivated by two main ideas:

1. That the brain provides a proof by example that intelligent behaviour is possible. From this perspective, a possible path to building real intelligence in apparatuses would be to reverse engineer the computational principles behind the brain, duplicating its functionality.
2. Deep learning models can help us to understand the brain and the principles tha underlie human intelligence. These models can shed light on fundamental scientific questions about how the brain works, besides their ability to solve engineering questions and help in artistic and creative ones.

Despite its limitations, by gathering and analysing knowledge from experience, machine learning algorithms do not require human operators to formally specify all the knowledge that the computer needs, as opposed to the GOFAI descriptive paradigm. Thus, it was easy to use an off the shelf algorithm to identify people and walls in any environment in Amigóide 2.0, making a special room for the automaton

[5] https://developer.nvidia.com/embedded/jetson-nano-developer-kit

unnecessary. Version 2.0 employs a cheap Raspberry Pi V2 Camera[6], replacing both the infrared and the colour sensors. Using NVIDIA VisionWorks API[7] in conjunction with OpenCV[8], it was feasible to have the 2.0 version working in less than one week, including the port of the original C code to Python. The level of flexibility provided by this new setup allowed version 2.0 to keep track of the part of the environment it cannot see, through the use of an internal state reflecting the unobserved aspects of the current state, making it a model-based reflex agent rather than its previous simple reflex agent condition.

As an aesthetic decision, the plastic black cylinder was kept as the main body, making both versions look practically the same. In terms of its inner-logic, the new outputs originated from the deep learning algorithms fed into a simplified GOFAI structure which is similar to the one deployed in version 1.0.

Amigóide's future looks bright, with new exhibitions already organized and a roadmap that has as its next step the implementation of an English speech recognition system, which will allow the automaton to hear and interact at a more sophisticated level with its human friends.

REFERENCE

AlphaGo (2017), Film. Director Greg Kohs.

Brooks, R. (1991) *Intelligence without Reason* https://people.csail.mit.edu/brooks/papers/AIM-1293.pdf (accessed 9th July 2019).

Brooks, R. (1987) Intelligence without representation. *Artificial intelligence*, *47*(1), 139-159.

Carey, J. (1983). Technology and Ideology: The Case of the Telegraph. *Prospects, 8*, 303-325.

Dreyfus, H. (1992). *What computers still can't do: a critique of artificial reason*. Cambridge: MIT Press.

Floridi, L. (1999). *Philosophy and computing: an introduction*. London: Routledge.

Ford, M. (2018). *Architects of intelligence: the truth about AI from the people building it.* Birmingham: Packt Publishing.

Gelertner, D. (1994) *The Muse in the Machine*. New York: Free Press.

Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep learning*. Cambridge: MIT Press.

Haugeland, J. (1993). *Artificial intelligence: the very idea*. Cambridge: MIT Press.

Marcus, G. (2018). *Deep Learning: A critical appraisal.* ArXiv: 1801.00631 (Jan. 2018). https://arxiv.org/abs/1801.00631

MIT (2019) Website. https://hcai.mit.edu (accessed 9th July 2019).

Mori, M., MacDorman, K., Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, no. 2: 98-100. https://doi.org/10.1109/MRA.2012.2192811

Plato (2007). *The republic*. London: Penguin.

Prosseda, R. (2014) https://www.robertoprosseda.com/media/projects/000008/attachments/download/teotronico-en-2014.pdf

Riek, L., Rabinowitch, T., Chakrabarti, B., Robinson, P. (2009) Empathizing with Robots: Fellow Feeling along the Anthropomorphic Spectrum. *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1-6. Amsterdam: IEEE. https://doi.org/10.1109/ACII.2009.5349423.

Russell, S., Norvig, P. (2016). *Artificial Intelligence: A modern approach*. Essex: Pearson.

Small, C. (1998) *Musicking*. Middletown: Wesleyan Press.

Vear, C. (2019) *The Digital Score*. New York: Routledge.

[6] https://www.raspberrypi.org/products/camera-module-v2/
[7] https://developer.nvidia.com/embedded/visionworks
[8] https://opencv.org/