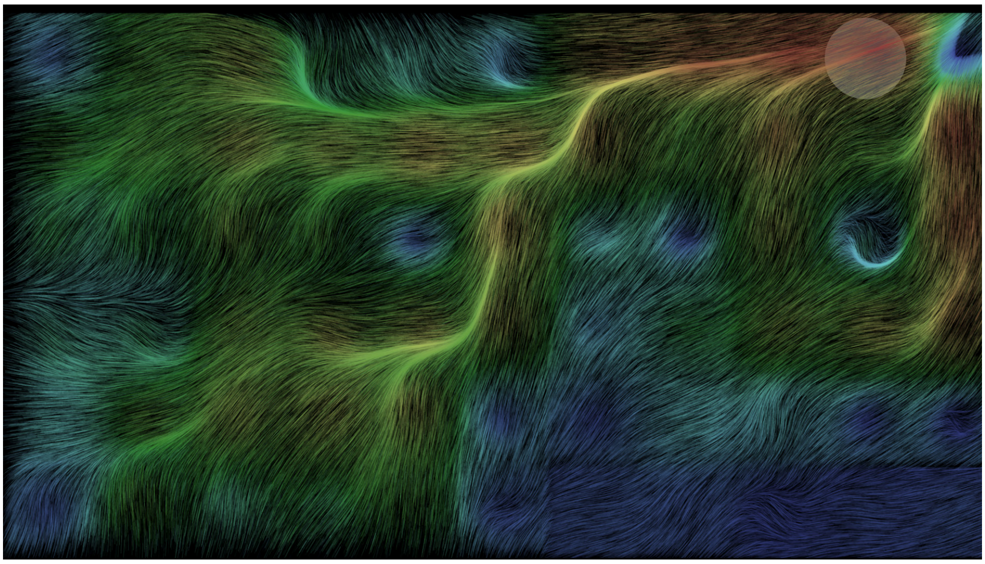
xCoAx 2023 11th Conference on
Computation, Communication,
Aesthetics & X
2023.xCoAx.org
Weimar, Germany

Greedy Agents and Interfering Humans



Tatsuo Unemi
unemi@soka.ac.jp
Soka University, Hachioji, Japan

Philippe Kocher
philippe.kocher@zhdk.ch
Zurich University of the Arts, Switzerland

Daniel Bisig
daniel.bisig@zhdk.ch
Zurich University of the Arts, Switzerland

DOI 10.34626/xcoax.2023.11th.363

This work addresses the coexistence of humans and an AI system by presenting an installation in which up to three visitors can interact with an agent that learns to move through an environment. The learning process, a central element of the installation, is made perceptible through visualisation and sonification. Since visitors can not only observe the advancement of the learning process but also actively support or impede it, the underlying simulation is not simply a black box but open to human perception and collaboration.

Keywords: Reinforcement Learning, Generative Art, Interaction, Audio-visual Installation.

Description

Greedy Agents and Interfering Humans is an interactive tabletop installation in which visitors can witness and interact with an agent that learns to move through an environment. The agent and its learning process are rendered perceivable by visualising and sonifying the states of the agent and its environment. The visuals are projected on a table surface, and the sounds can be listened to on headphones (Fig. 1). By placing their hands on the table surface, visitors can affect the agent and the state of its environment. This interaction also influences — yet only indirectly, mediated by the simulation — the visuals and sounds that the visitors eventually perceive (Fig. 2).

Figure 1: Schematic installation setup.

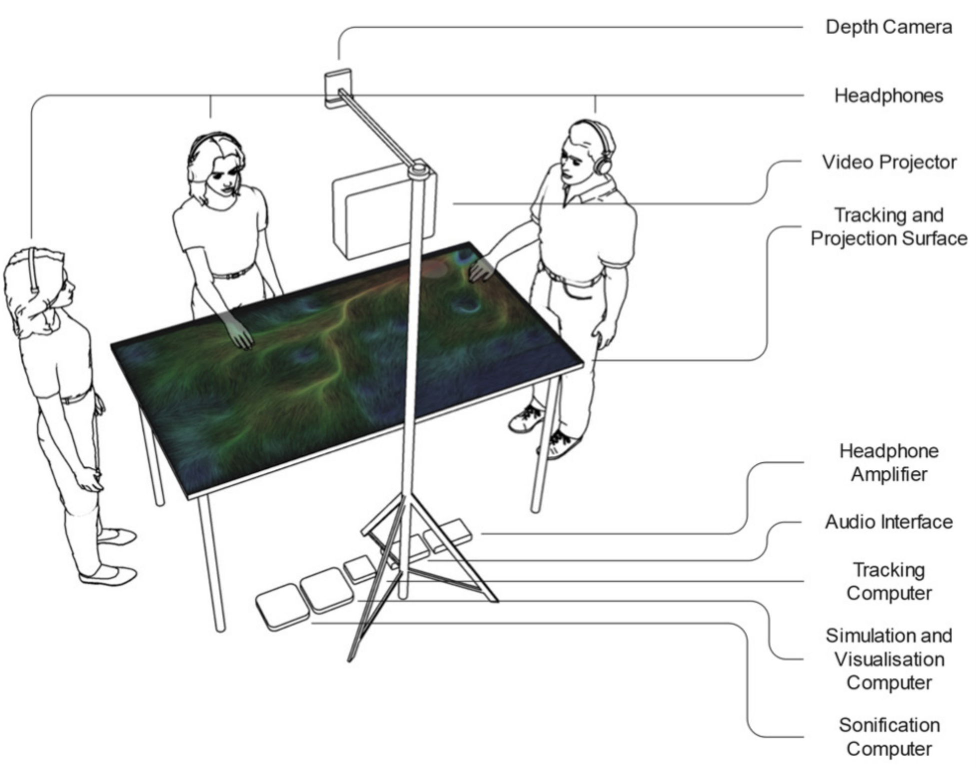
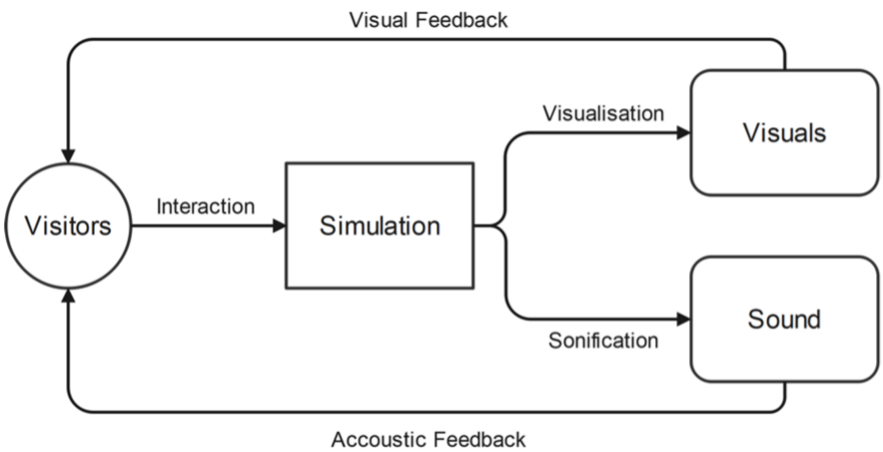


Figure 2: Flowchart of the interaction.



In the broadest sense, this installation addresses the coexistence between humans and AI systems. Rather than being in a user-tool relationship, both jointly inhabit an environment that is at the same time real and virtual. In the same vein, the authors understand creativity as a joint effort of several human and non-human actors connected in a network. This installation continues the authors' earlier works dealing with the coexistence of human and non-human actors (Bisig and Kocher 2015; Bisig and Unemi 2011; Bisig and Unemi 2010; Bisig and Unemi 2009).

The process at the heart of the installation is reinforcement learning. This learning paradigm has been researched for more than one hundred years in the fields of psychology and ethology (Thorndike 1898; Skinner 1953; Sutton 1998). It is a framework to explain animal/human learning mechanisms based on the modification of behaviour through experience. By visualising and sonifying this learning process and by exposing it to user interaction, it is rendered perceivable. Instead of being a black box, the learning process unfolds before the visitors' eyes and becomes part of the aesthetic expression of an artwork.

Technical Details

Simulation

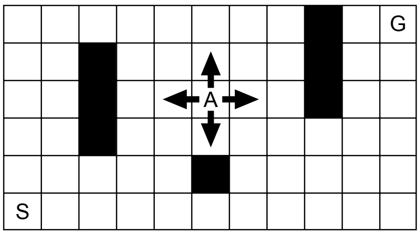


Figure 3: The 6 x 11 simulation grid. S: start, G: goal, A: agent, black squares: obstacles.

The simulation on which this installation is based is an implementation of a Q-learning algorithm (Watkins 1989). It employs a simple navigation task: an agent has to find the shortest path from a start to a goal location. The environment in which the agent moves about consists of a grid of 6 x 11 squares which might also include a certain number of obstacles (Fig. 3). In the beginning, when it starts learning, the agent does not know anything about the world and moves randomly from square to square. Upon reaching the goal location, it receives a reward depending on the efficiency of its search. Then it is put back to the start square, and the search begins again. During the process of learning, the agent builds a memory of how fruitful it is on each square to move on in a specific direction (up, down, left, right) to obtain the highest possible reward. The learning process continues until the number of simulation steps or the number of times the goal was reached exceeds a predefined value. In that case, the memory is reset and the whole learning simulation restarted.

Interaction

Up to three visitors can interact at the same time with the installation. The interaction is based on tracking the positions of the visitors' hands on top of the table surface on which the visualisation of the simulation is projected. The hands are detected by means of a distance camera mounted above the table and pointing vertically down. The positions are derived from the front-most points of the hands' contours and mapped to the square of the simulation world with which the position overlaps.

By interacting, the visitors influence the agent's learning process through various effects, such as:

- the hand guides or blocks the agent's movements.
- the hand traces future paths for the agent.
- the hand creates obstacles for the agent.

Visualisation

The simulation's state is visualised and projected on the table. The agent itself is shown as a white circle. The agent's memory is interpreted as a vector field representing the preferred direction at each square calculated as the sum of the four vectors corresponding to the directions in which the agent can move. A particle flow animation visualises this vector field by some hundred thousand short line segments moving across the table, pushed by the forces of the vectors. Each line segment is drawn in a colour indicating its moving speed. The visitors recognise through this visualisation the progress of the agent's learning. While the particle flow is not yet pronounced at an early stage of learning (Fig. 4), as the simulation advances, it becomes more clearly directed towards the goal (Fig. 5).

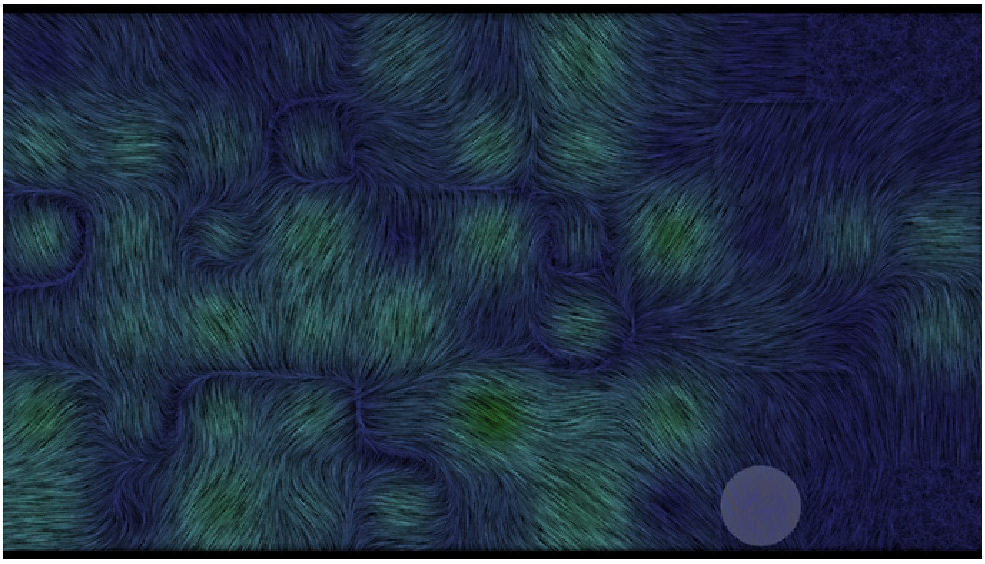


Figure 4: The visualisation of the simulation at an early stage of learning.

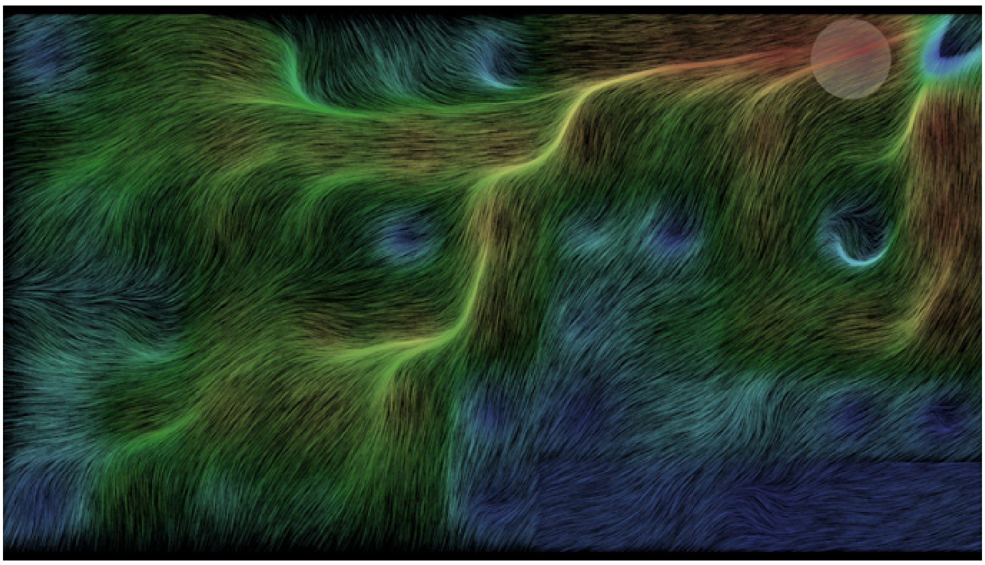


Figure 5: The visualisation of the simulation at a later stage of learning.

Sonification

The sound played back to the visitors via headphones reflects the state and the dynamics of the simulation. By moving their hands across the table, the visitors can acoustically scan the landscape as the current location of the hand determines which part of it is made audible. The properties of each square, i.e. the values of its vectors, the existence of an obstacle or the momentary presence of the agent, are taken as parameters for a real-time sound synthesis algorithm. Unlike the visualisation, which reflects the current state of the simulation, the sonification algorithm also creates a history as it keeps track of previous values. Even if the learning simulation restarts from time to time, it leaves traces in the sonic environment, and the acoustic output will evolve and never be exactly the same.

References

Bisig, Daniel and Philippe Kocher. 2015. "DRIFT – Virtual Sand in Augmented Space." *Proceedings of the 18th Generative Art Conference*, Venice, Italy, pp. 51–64.

Bisig, Daniel and Tatsuo Unemi. 2009. "Swarms on Stage – Swarm Simulations for Dance Performance." *Proceedings of the 12th Generative Art Conference,* Milano, Italy, pp. 105–114.

Bisig, Daniel and Tatsuo Unemi. 2010. Cycles-blending natural and artificial properties in a generative artwork. *Proceedings of the 13th Generative Art Conference*, Milan, Italy, pp. 140–154.

Bisig, Daniel and Tatsuo Unemi. 2011. "From Shared Presence to Hybrid Identity." *Proceedings of the Consciousness Reframed Conference,* Lisbon, Portugal, pp. 48–53.

Skinner, B. F. 1953. *Science and Human Behavior*. New York: MacMillan.

Sutton, Richard S. and Andrew G. Barto. 1998 *Reinforcement Learning: An Introduction* (2nd ed. 2018). Cambridge, Massachusetts: MIT Press.

Thorndike, Edward L. 1898 (1911). "Animal Intelligence: an Experimental Study of the Associative Processes." *The Psychological Review: Monographs Supplements,* 2(4), i–109.

Watkins, Christopher. J. C. H. 1989. *Learning from Delayed Rewards*. Ph.D. Thesis. University of Cambridge.