Research Statement | Thomas Langerak

thomas.langerak@inf.ethz.ch | thomaslangerak.nl

Research Vision

The goal of my research is to make **the interaction between intelligent systems (e.g., robots, semi-autonomous vehicles, and digital assistants) and humans more intuitive**. I believe that Contextual and Embodied Artificial Intelligence (AI) causes the boundaries between the physical and digital worlds to blur, necessitating technological advances that **prioritize human-centric control of intelligent systems**. My research aims to enable a future where all humans can interact with intelligent systems instinctively, unobtrusively, and with minimal effort required to master.

I believe that the current state of systems is reactive: users provide the system with an explicit input, and the system reacts to that input. Contextual AI enables intelligent reactions; the system combines explicit user input with implicit contextual understanding. This contextual understanding is necessary for Embodied AI, which allows intelligent systems to operate in the real world. Yet, **systems still require explicit user commands**, thereby making interactions unnatural, cumbersome, and unsatisfactory. In an ideal world, the intelligent system would not solely rely on explicit instructions from the user but could infer the user's desires from their behavior. In other words, we need to **transform systems from reactive to proactive**. I believe that **embedding human behavior models in optimal control strategies is crucial** for this transition.

To achieve this goal, I create novel **computational approaches** that enable a more intuitive interaction, study and explore how these technologies are perceived and used, and apply these insights to inform the design of novel methods. I believe that exploiting the notion of **humans as rational, goal-driven beings**, will allow us to create intelligent systems that enable implicit, thus intuitive, interactions.

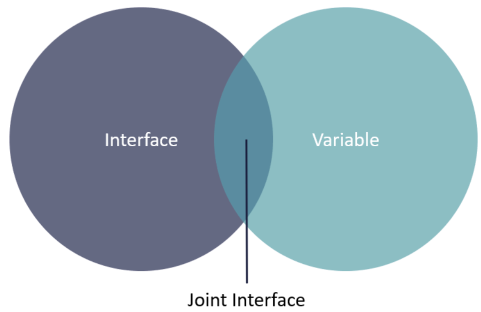
Research Contributions

Viewing a system as a tool, we traditionally used such tools to directly manipulate variables, like using a hammer to drive a nail. Recently, this approach has evolved: we now interact with physical systems that communicate with intelligent agents to manipulate variables for us, such as controlling a Nest thermostat to adjust the temperature. **With the rise of Contextual and Embodied AI, both the user and the agent interact directly with the variable**, making it both the target of manipulation and the interface for interaction, exemplified by users interacting with code through Copilot. This convergence of variables and interfaces introduces **challenges in rethinking interface design and balancing user autonomy with system automation.**

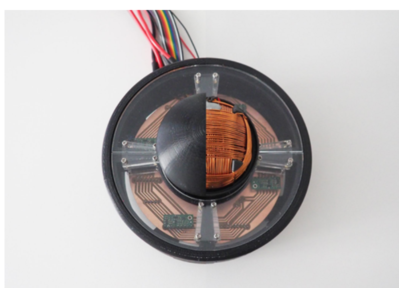
My research follows these two connected threads of work: 1) the design, creation and evaluation of **joint interfaces**, interfaces on which both a human and artificial agent interact; and 2) the exploration of **cooperative control strategies** that combine **user models with optimal control strategies.**

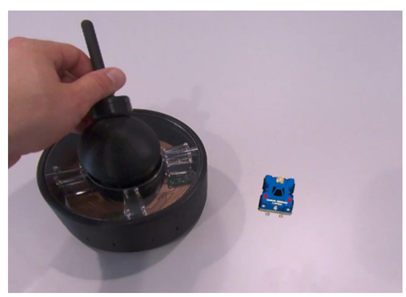
Interface + Variable = Joint Interface

Traditionally, **an interface is defined as a shared boundary for information exchange between separate computer system components**. A joint interface retains these properties but also **serves as a variable that the user seeks to control (right),** such as a piece of code. I am not arguing that all variables are interfaces or vice versa. I am saying that these historically separate elements have moved closer over time, due to contextual AI, and that there is now a partial overlap.



I have explored this overlap in three projects; in which I specifically focused on physical interfaces. In all my projects the intelligent agent and the user act (using force) on the same style of input device (a pen).



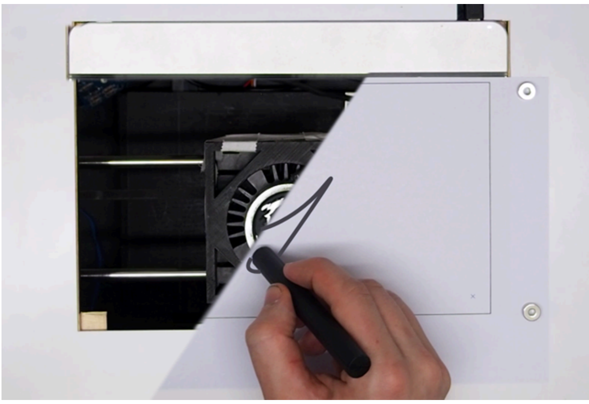


Omni [Langerak2020a] (above) is an haptic device. A symmetric electromagnet is used in combination with a dipole magnet model and a simple control law to deliver dynamically adjustable forces onto a hand-held tool. The tool only requires an embedded permanent magnet and thus can be entirely untethered. Here the haptic feedback is elicited on the tool, while the user interacts with the tool. Omni showed that **physical joint interfaces provide more intuitive and accurate user inputs.**

[Langerak2020b] extends Omni to include sensing of the tool position via the magnetic field. Where Omni required external tracking devices; the spatial haptic capabilities of Omni 2.0 are enabled by a novel gradient-based method to reconstruct the 3D position of the permanent magnet in midair using the measurements from eight off-the-shelf hall sensors that are integrated into the base. In [Langerak2022] we extend this to a deep learning method that improves tracking latency, frequency, and accuracy. Omni 2.0 teaches us that by **incorporating advanced sensing technologies directly within our tools, we can significantly reduce reliance on external devices, paving the way for more seamless and integrated human-computer interactions.**

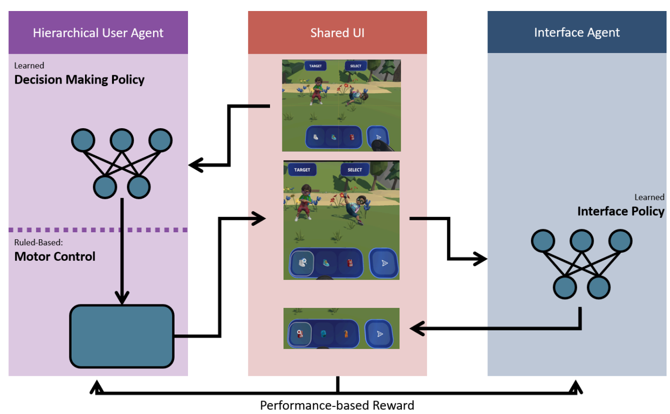
Cooperative Control Strategies

The contextualization and embodiment of AI have a second implication for interfaces and variables: **rather than users having full ownership over the interface, ownership is now shared**. For instance, until recently, users would write code on their own. However, nowadays, both the user and CoPilot can write code together, simultaneously. This raises questions about who can edit and manipulate the code and to what extent, at any given time. **We need to balance user autonomy with system automation.**



One of the core limitations of the Omni iterations is that the magnetic field rapidly declines over distance, thereby limiting the working area. In [Langerak2020c], we overcame this by mounting a cylindrical electromagnet on a biaxial linear stage (left). However, this approach complicates the control dynamics. If the magnetic actuator is too close and active, the feedback perceived by the user is too strong, causing the user to lose any sense of autonomy. However, if the magnet is too far away, no haptic feedback is perceived, rendering the system useless. **In our work, we demonstrate that by incorporating user and world models into a control strategy, we can guide the user while still preserving their autonomy.**

We overcome the core limitations of having to provide user and world dynamics in [Langerak2023] (right). **Our control policy is learned by playing a multi-agent stochastic game through reinforcement learning.** One agent learns to interact with an interface, mimicking a user, while another learns the underlying task structures and user behavior by observing states and actions. This project **shows that control strategies can be learned without relying on hand-crafted heuristics.**



Future Research Agenda

In my future research, I plan to focus on three distinct aspects and opportunities to realize the vision of human-centric control over intelligent systems.

Human Sensing & Inference

I believe that the impending transformation of the interface will **integrate the world at large—including the users themselves—into the interface itself**. For instance, consider a scenario in which a robot and a human collaborate to clean an apartment. The user should **not need to explicitly instruct the robot** on what to clean; instead, the robot should be capable of inferring this from the state of the world and the user's actions.

This requires two main components. Firstly, **accurate human state estimation** is essential. This includes **not only the physical** state in the world, which can be captured through computer vision and other modalities, but **also latent states** such as expertise, fatigue, and intent. Secondly, **accurate sensing of the world** state and reasoning is required. The system should be aware of its environment and capable of predicting future states from the current context, such as understanding that an apple will fall if released.

Human Behavioral Models

However, sensing the human state is insufficient. To apply current state estimations to intelligent control, we need to **predict future human—both physical and latent—states.** This requires behavior models that capture the complex dynamics of human behavior.

To solve this problem, we need **advances in data-driven Reinforcement Learning, such as imitation learning (IL) and inverse reinforcement learning (IRL), combined with existing cognitive models.** By integrating the strengths of these three approaches, we aim to develop methods that approximate human strategies, are generalizable, and computationally feasible. These models will not only enhance our understanding of human behavior but also improve how we design interfaces and interaction paradigms between humans and machines.

Cooperative Control

Finally, we need to **integrate the human sensing, inference, and behavioral models into intelligent control strategies**, paving the way for more intuitive and seamless interactions between humans and machines that respect and enhance human decision-making processes.

I foresee a future in **hierarchical and multi-agent reinforcement learning**, where predictive human models can be efficiently integrated into control strategies. Exactly how this integration will occur remains an open question. The development of such systems promises to significantly advance our ability to design intelligent systems that truly augment human abilities, marking a significant milestone in human-computer interaction.

References

[Langerak2023] MARLUI: Multi-Agent Reinforcement Learning for Goal-Agnostic Adaptive UIs. Thomas Langerak, Sammy Christen, Mert Albaba, and Christoph Gebhardt, Otmar Hilliges. *Under Review*

[Langerak2022] Robust Real-Time Tracking of Axis-Symmetric Magnets via Neural Networks. Mengfan Wu*, Thomas Langerak*, Otmar Hilliges, and Juan Zarate. *Under Review*

[Langerak2020a] Contact-free nonplanar haptics with a spherical electromagnet. Juan Jose Zarate*, Thomas Langerak*, Bernhard Thomaszewski, and Otmar Hilliges. *In 2020 IEEE Haptics Symposium (HAPTICS), 2020*

[Langerak2020b] Omni: Volumetric sensing and actuation of passive magnetic tools for dynamic haptic feedback. Thomas Langerak*, Juan José Zárate*, David Lindlbauer, and Christian Holz, Otmar Hilliges. *In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, 2020*

[Langerak2020c] Optimal control for electromagnetic haptic guidance systems. Thomas Langerak, Juan José Zárate, Velko Vechev, David Lindlbauer, Daniele Panozzo, and Otmar Hilliges. *In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, 2020*