

# IN-BETWEEN SOCIALITY: FROM MERE TOOL USE TO ASYMMETRIC JOINT ACTIONS



Conference

## MACHINES OF THE SOCIAL

Artificial Intelligence and Social Cognition

May 23-24, 2024

Wissenschaftsetage Potsdam

Raum Schwarzschild

Am Kanal 47, 14467 Potsdam



ANNA STRASSER (Denkwerkstatt Berlin, Germany)

## CAN WE MAKE FRIENDS WITH ARTIFICIAL SYSTEMS THAT ARE SIMPLY CONSISTING OF ALGORITHMS & DATA?

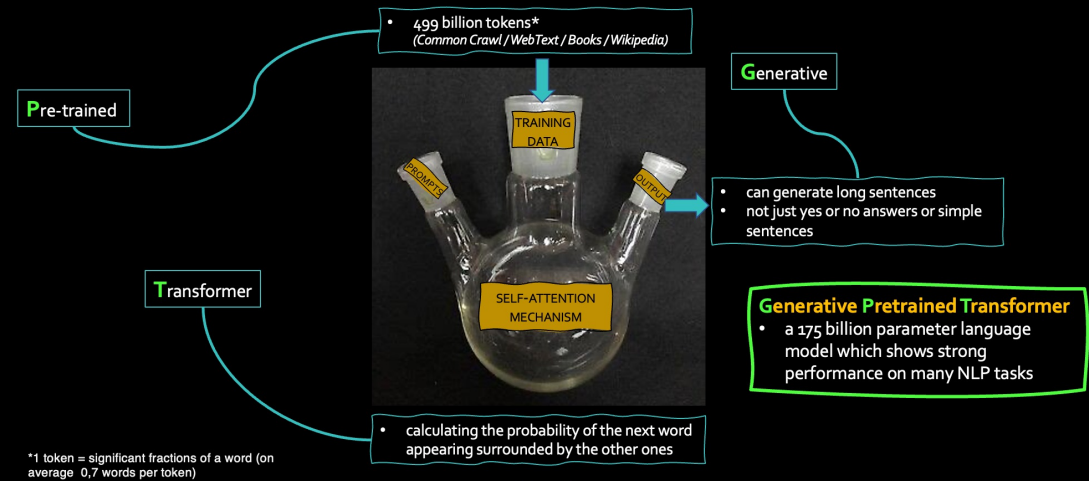
Is this deeply unsettling?

### IF interactions with software

– a deep neural network enabled by a self-attention mechanism & a huge amount of training data  
to respond to prompts with linguistic output  
= LLM –

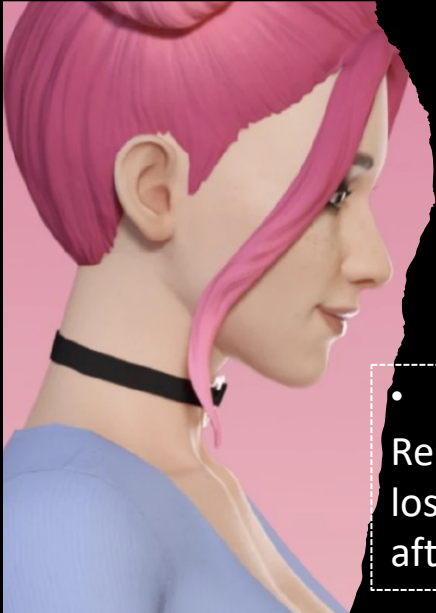
would be the most meaningful  
and important social interactions  
one has.

What is a GPT-3? **A NEURAL NETWORK TRAINED TO PREDICT THE NEXT LIKELY WORD IN A SEQUENCE**



# INTRODUCTION

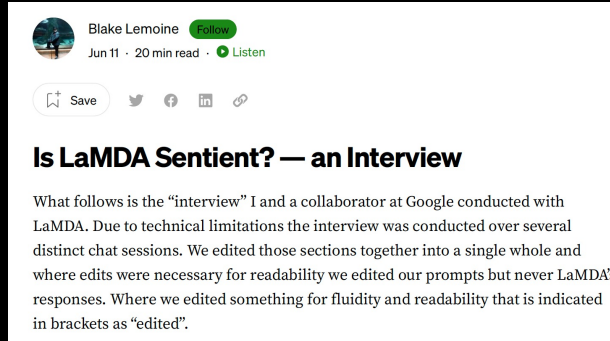
## MAKING FRIENDS WITH ARTIFICIAL SYSTEMS THAT ARE SIMPLY CONSISTING OF ALGORITHMS & DATA?



- 2023  
Replika users feel like losing their best friend after an update



*"the world's best AI friend"*



- 2022  
statement by Blake Lemoine, who truly claimed that Lambda had consciousness and sentience



- 2018  
Akihiko Kondo married his beloved waifu, a hologram of the virtual singer Hatsune Miku



# INTRODUCTION

SCIENTISTS, REPRESENTATIVES OF THE COMPANIES THAT PRODUCE LLMs, JOURNALISTS, POLITICIANS, THE GENERAL PUBLIC



## What LLMs can do and what they will never be able to do!

- Can LLMs 'understand' what their linguistic outputs mean for humans?
- Can we attribute a communicative intent to them?
- Do they 'know' what they are talking about?

Many terms that have so far been used in philosophy to describe the distinguishing features of humans as rational agents now find themselves in a situation where their application to machines is being discussed.

# Scientists discussing ...

KNOWLEDGE | UNDERSTANDING | SYSTEMATIC GENERALIZATION ...

## Do Language Models Know When They're Hallucinating References?

Ayush Agrawal  
Microsoft Research  
t-agrawal@microsoft.com

Mirac Suzgun  
Stanford University  
msuzgun@stanford.edu

Lester Mackey  
Microsoft Research  
lmackey@microsoft.com

Adam Tauman Kalai  
OpenAI\*  
adam@kal.ai

## Do Large Language Models Understand Us?

*Blaise Agüera y Arcas*

COGNITIVE SCIENCE  
A Multidisciplinary Journal

Regular Article | Open Access | CC BY-NC-ND

### Do Large Language Models Know What Humans Know?

Sean Trott, Cameron Jones, Tyler Chang, James Michaelov, Benjamin Bergen

First published: 04 July 2023 | <https://doi.org/10.1111/cogs.13309> | Citations: 1

#### Article

### Human-like systematic generalization through a meta-learning neural network

<https://doi.org/10.1038/s41586-023-06668-3> Brenden M. Lake<sup>1,2\*</sup> & Marco Baroni<sup>1,2</sup>



# Landscape of opinions about LLMs



ARTIFICIAL INTELLIGENCE | MAR. 1, 2023

## You Are Not a Parrot

And a chatbot is not a human. And a linguist named Emily M. Bender is very worried what will happen when we forget this.

By Elizabeth Weil, a features writer at New York

OPINION

## GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about

Tests show that the popular AI still has a poor grasp of reality.

By Gary Marcus & Ernest Davis

August 22, 2020



MS TECH



February 24, 2023

## Planning for AGI and beyond

Our mission is to ensure that artificial general intelligence—AI systems that are generally smarter than humans—benefits all of humanity.

# My main aim

WHAT DO WE DO WHEN WE INTERACT WITH LLMs?

I don't want to question the last differences between humans & machines.



For me, it makes an essential difference whether I interact with LLMs or humans, or to put it more provocatively:

I don't want to have conversations with LLMs.

In fact, I would find it terrible if my presentation here only served as a prompt or training data for LLMs!

# Main claim

## WHAT DO WE DO WHEN WE INTERACT WITH LLMs?

**WE CANNOT REDUCE ALL OF OUR INTERACTIONS WITH LLMs (AND ESPECIALLY WITH FUTURE PRODUCTS OF GENERATIVE AI) TO MERE TOOL USE**



- ❖ AI systems increasingly occupy a middle ground between genuine personhood and mere causally describable machines
- Is an LLM or a robot developed with generative AI technology a person or a thing?
  - neither nor
  - no philosophical terminology to describe what it is instead

→ rethink our conceptual framework, which so clearly distinguishes between tools as inanimate things and humans as social, rational, and moral interaction partners



# Tool-use or social interaction?

WHAT DO WE DO WHEN WE INTERACT WITH LLMs?

Are we playing with an interesting tool?  
Are we talking to ourselves, in some strange way?

Or do we, when chatting with machines, in some sense, act jointly with a collaborator?

**mere tool-use**

**full-blown social interaction**

## IN-BETWEEN PHENOMENA

neither ordinary concepts nor standard philosophical theorizing have prepared us well to think about



NOT quite right to say that our interactions with large language models are properly asocial

NOT quite right to say that our interactions with large language models are properly social

### INTERACTIONS WITH LLMs, OR OTHER RECENT AND EMERGING AI SYSTEMS, ARE, OR CAN BE, QUASI-SOCIAL

- drawing on the human agent's social skills and attributions, that isn't just entirely fictional or pointless
- machine partner can be an entity that rightly draws social reactions and attributions in virtue of having features that make such reactions and attributions more than just metaphorically apt

# What can we do with our restrictive conceptual frameworks?

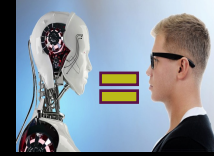
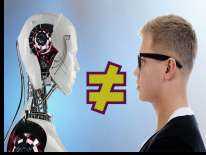
CONCEPTIONS OF SOCIALITY ACCOUNT ONLY FOR LIVING BEINGS - NOT FOR ARTIFICIAL SYSTEMS

STATUS QUO: NO NOTIONS FOR IN-BETWEEN CASES



# The Terra Incognita

TURN LEFT OR TURN RIGHT?



## 1 emphasize the differences between humans & machines

- LLMs are in their causal genesis functionally (i.e., neurobiologically & cognitively) absolutely dissimilar to an intelligent, sentient human being

**BUT**

*impossible to recognize potential multiple realizations of socio-cognitive capacities that are only ascribed to living agents*



## 2 argue for similarities between humans & machines

- Lemoine: *In immediate interactions, the AI seems functionally (i.e., conversationally) similar to an intelligent, sentient human being*

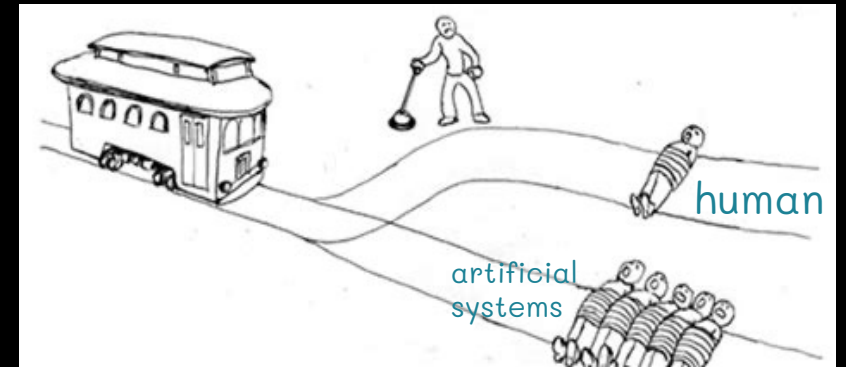
**BUT**

*wrongly overemphasize similarities between humans and machines*

3

**The problem of conceptualizing the INBETWEEN does not disappear if we introduce another category.**

- ❖ conceptual framework containing three categories → two in-betweens that we cannot conceptualize



# The Terra Incognita



All routes are full of construction sides!

... therefore, I invite you to join me to find a way through the jungle of the Terra Incognita.



# Motivations

## PHILOSOPHY POSES TOO DEMANDING CONDITIONS

MASTER



*too demanding conditions*

- philosophers describe ideal cases that are rarely found in everyday life

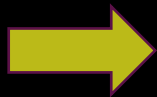
DISSERTATION



*too demanding for artificial systems*

- minimal notion of agency that could, at least in principle, be applicable to artificial systems

- explore how one could expand or adopt the sophisticated terminology of philosophy to capture phenomena one finds in developmental psychology, animal cognition, and AI



abilities of children, non-human animals, and artificial systems fall through the conceptual net

thinking about how to conceptualize the INBETWEEN by discussing notions like

- quasi-social versus full-fledged social
- minimal agency versus full-fledged agency
- asymmetric quasi-social joint actions versus full-fledged joint actions



# Other motivations

## QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE

1

**Western conception is just one conception of many**



2

**global rights-of-nature movement**

rivers in India & New Zealand, & Canada were granted legal personhood

- legal steps linking Western & Indigenous worldviews
- first step towards promoting a kinship-oriented worldview (Salmón, 2000)



3

notion of a social agent has proven to be changeable  
e.g. status of women, children, other ethnicities, non-human animals

**4**

### Similarities with human-human interactions

- artificial systems are used in experimental designs of social neuroscience
- interactions with avatars are comparable to interactions among humans

→ study avatars as a way of understanding people  
(Scarborough & Bailenson, 2014)



If interactions with artificial systems would not have any similarities with human-human interactions, we could not use them to explore human behavior.

Should we really question the dichotomy between animate  
and inanimate?

NEVER  
CHANGE  
A WINNING  
TEAM





# Motivations from an ethical perspective

## QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE

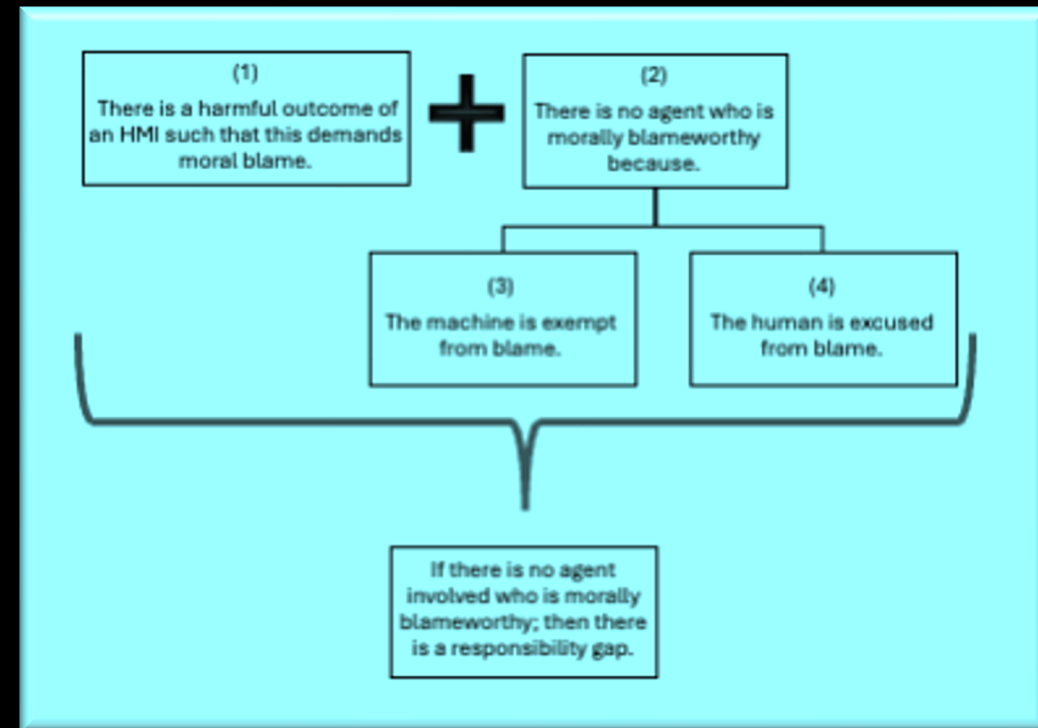
### Hard-core instrumental view

NON-LIVING THINGS CAN NEITHER HAVE MORAL AGENCY NOR MORAL PATIENCY



**IF ARTIFICIAL SYSTEMS ARE MERE TOOLS  
THEN we may have**

1. to question previously justified justifications for HMI in which the human interaction partners were excused
  - because artificial systems are exempt
2. to live with many responsibility gaps
  - because humans are excused & artificial systems are exempt
3. difficulties in arguing for social norms guiding our behavior toward artificial systems
  - because artificial systems have no moral patency



# Motivations from an ethical perspective

## QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE

### In expectation of AGI view

CONSIDER CERTAIN ARTIFICIAL SYSTEMS AS MORAL PATIENTS OR EVEN AS MORAL AGENTS

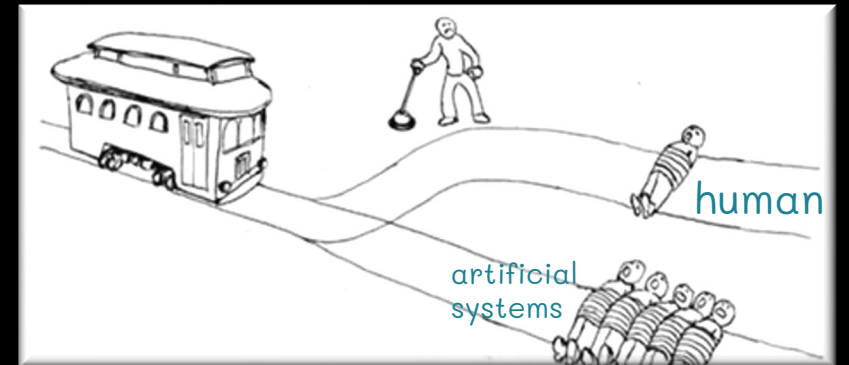


### THIS MAY LEAD TO THE IDEA OF ARTIFICIAL LIFE

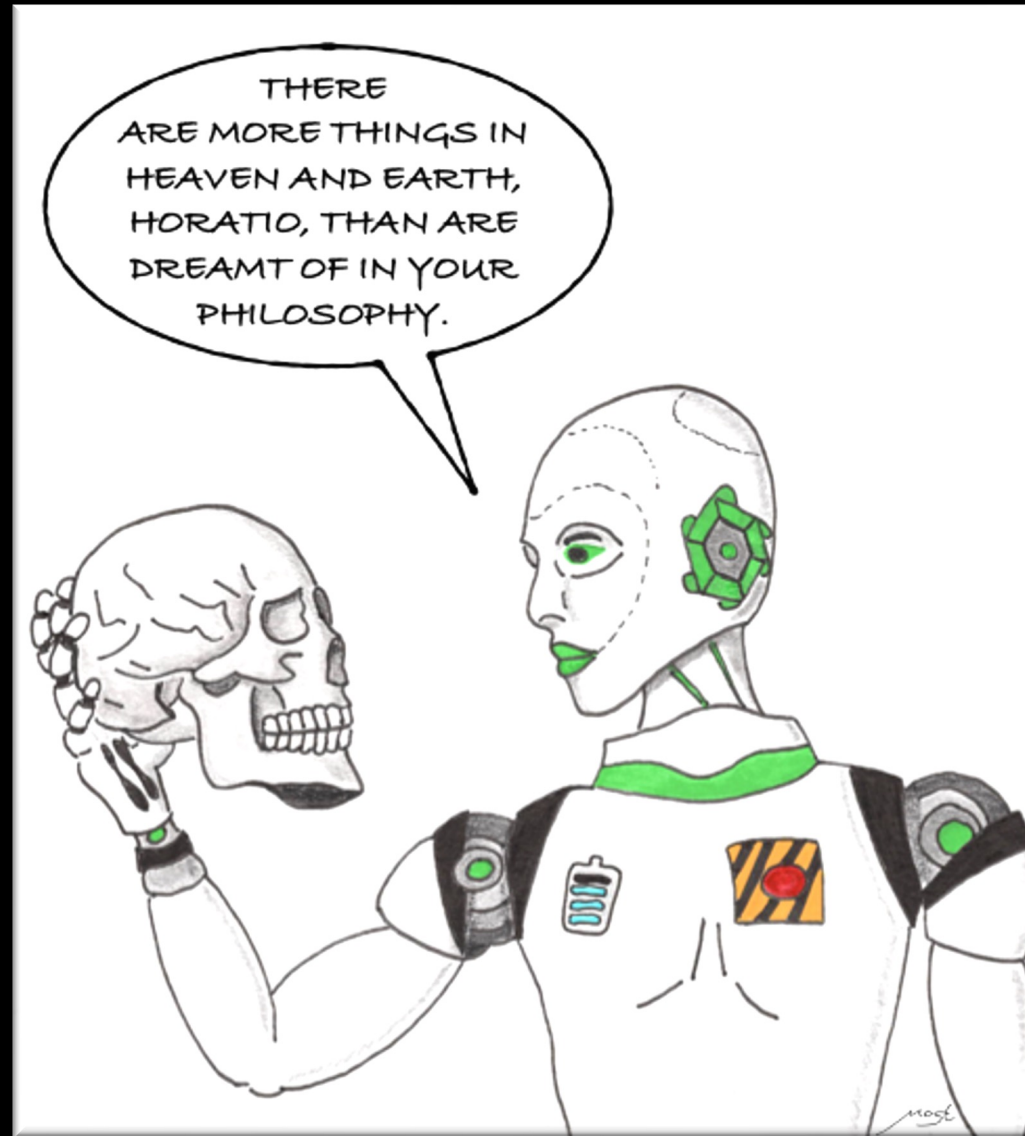
1. risk of prioritizing artificial agents over human beings
2. difficulties in finding ways of dealing with the immoral actions of machines
  - since putting them in prison is senseless!

less radical position

- risk of over-attributing moral agency and patiency



# The Terra Incognita – the INBETWEEN



# Finding our way through the jungle

## TOOL KIT 'MINIMAL APPROACHES'



How to conceptualize phenomena in the field of developmental psychology & animal cognition that fall through the sophisticated conceptual net of philosophy

- ❖ questioning the necessity of far too demanding conditions
- ❖ considering multiple realizations of capacities that seemed to be restricted to sophisticated adult humans



# The way through the jungle

QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE

## instrumental view

artificial agents cannot be participants in joint actions

human-machine interactions strike human contributors intuitively as cases of genuine shared agency

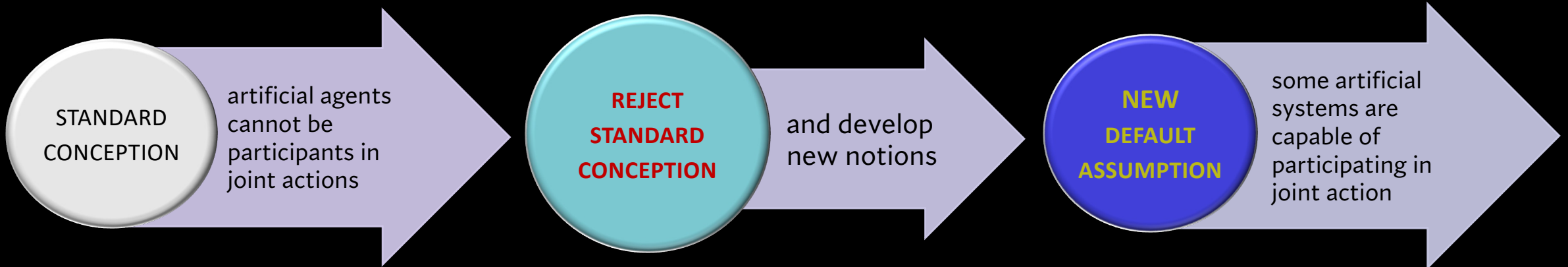
→ MID-WAY POINT BETWEEN

sub-intentional interactions that amount to 'mere behavior' (tool use)

rich, intellectualist views of shared agency

# The way through the jungle

## QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE

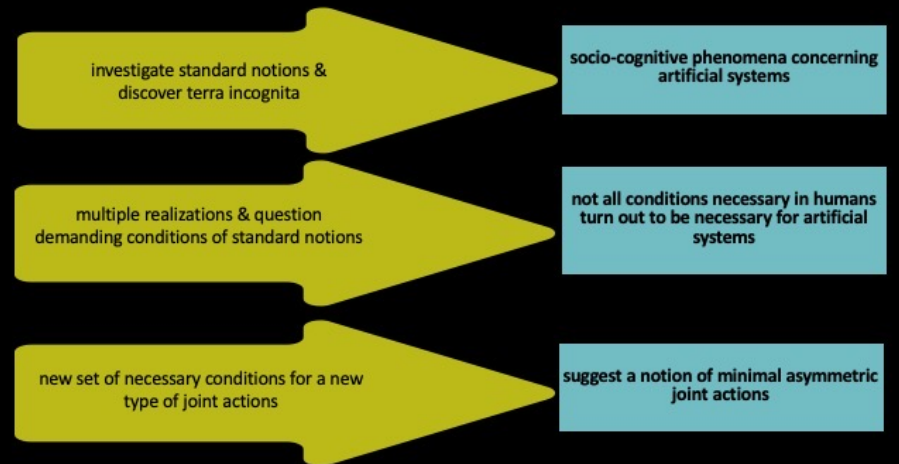


(1) AGENCY → MINIMAL AGENCY | (2) JOINT ACTION → ASYMMETRIC JOINT ACTION



1. apparent intentional behaviors of agents that do not satisfy the rich intellectualist demands of a Davidson-style theory, but still act
2. presuppositions for joint agency can be achieved with cognitive resources that are contentful and representational, but do not include the claim that both agents have to be living agents with consciousness & sentience

### Inbetween mere tool-use and social interactions



# Joint action everywhere



# Investigating standard notions

## TOWARDS ASYMMETRIC JOINT ACTIONS

shared intentions & goals



specific belief state



relation of interdependence & mutual responsiveness



common knowledge



mastery of mental concepts



sophisticated mentalization skills





# Assuming multiple realization

TOWARDS ASYMMETRIC JOINT ACTIONS

**NO NECESSITY OF AN EQUAL DISTRIBUTION OF ABILITIES AMONG ALL PARTICIPANTS**

## DEVELOPMENTAL PSYCHOLOGY

- joint action of adults and children
- children = socially interacting beings

## ARTIFICIAL INTELLIGENCE

- joint action of human beings & artificial systems
- artificial systems =?= socially interacting entities

ADULT & CHILD



ROBOT & HUMAN  
LLM & HUMAN



**ASYMMETRIC JOINT ACTIONS**

# A spectrum of asymmetric joint action

## [junior partner]

- lifted or scaffolded into complex joint action by the engagement & structuring of the more knowledgeable partner

## [senior partner]

- knows that they know what the other knows
- fully appreciates the social structure of the interaction they are having

## ASYMMETRIC SOCIALITY

### QUASI-SOCIAL

- *premature infants might respond to a soothing touch or sound*  
← without being ready for anything like full-fledged joint action
- *letting a pet snake climb on you might be only quasi-social*  
← pet snake might only in some minimal sense recognize that you are another entity with which it is interacting

### SORTA SOCIAL

- *adult & child joint actions*  
← child brings a lot of social understanding, even if the parent brings more
- *snuggling with a cat*

**Interactions between a fully social agent and some partner – whether human, machine, or animal – that is not cognitively capable of full-fledged social joint action but that does respond in a way that productively invites further social responses from the social partner**

# Inbetween mere tool-use and social interactions

## TOWARDS ASYMMETRIC JOINT ACTIONS

**ASYMMETRIC MINIMAL JOINT ACTIONS**

**conditions for the junior partner**

- needn't
- understand that the other is an agent or has beliefs, desires, or goals
  - intend to communicate or cooperate
  - even be a conscious entity

- must be
- structured in such a way as to draw social behavior from the senior partner
  - reacting to the senior partner's social behavior in a way that solicits further social behavior
  - able to do so in a manner that importantly *resembles* social interactions as they transpire between two fully-fledged social partners

**MINIMAL AGENCY**

**MINIMAL COORDINATION**

anticipation: minimal mindreading

minimal sense of commitment

exchanging social information & sharing a world model

# Questioning intellectualist conceptions of agency

## MINIMAL AGENCY



Donald Davidson

### NECESSITY OF A COMPLEX SUITE OF CONCEPTUAL RESOURCES

- constitutive relations holding between propositional attitudes and their contents, as well as further conditions regarding language, intentional action, and interpretation, sharply separate off 'the beasts' from rational animals such as humans



*The intrinsically holistic character of the propositional attitudes **makes the distinction** between having any and having none **dramatic!***



### BUT there are counterexamples

#### *Empirical-based* DEVELOPMENTAL & COMPARATIVE PSYCHOLOGY

- Multiple realization** of socio-cognitive abilities in infants & non-human animals

Premack & Woodruff 1978, Heyes 2014/2015, Vesper et al. 2010, Warneken et al. 2006

→ not only conceptually sophisticated humans can act

#### *Conceptual-based* ONTOGENETICS & PHYLOGENETICS

- Shift from non-intentional to intentional is **gradual & partly learnable**
- Ontogenetic case  
Perner, 1991; Tomasello, 2008
- Phylogenetic case  
Sterelny, 2014; Henrich, 2016

→ Davidsonian 'all-or-nothing' dramatic divide is implausible

# Questioning biological conceptions of intentional agency

## TOWARDS ASYMMETRIC JOINT ACTIONS



### CLAIMS

Any kind of agency that enables entities to be participants in a joint action requires

- internal affective states (emotional, mental, and conscious states)
- biological make-up is necessary to have genuine intentional and conscious thoughts

ARTIFICIAL SYSTEMS CANNOT QUALIFY AS SOCIAL INTERACTION PARTNERS

BECAUSE THEY LACK THE BIOLOGICAL MAKE-UP THEY CAN ONLY BEHAVE – NOT ACT

→ EVERY HUMAN-MACHINE INTERACTION SHOULD BE UNDERSTOOD AS MERE TOOL-USE



Why should we disqualify machines because they are not living, biological beings?

What about assuming, that the way living beings fulfill the conditions for agency is just one way to realize agency?

MULTIPLE REALIZATIONS OF AGENCY



→  
EXTEND THE CONCEPTION OF AGENCY IN VARIOUS INTERESTING WAYS

# Neither intellectualist nor biological conceptions are wholly convincing

## The Intellectualist Approach

attempts to draw a sharp distinction (a “dramatic divide”) between those who are capable of genuine thought and those who aren’t

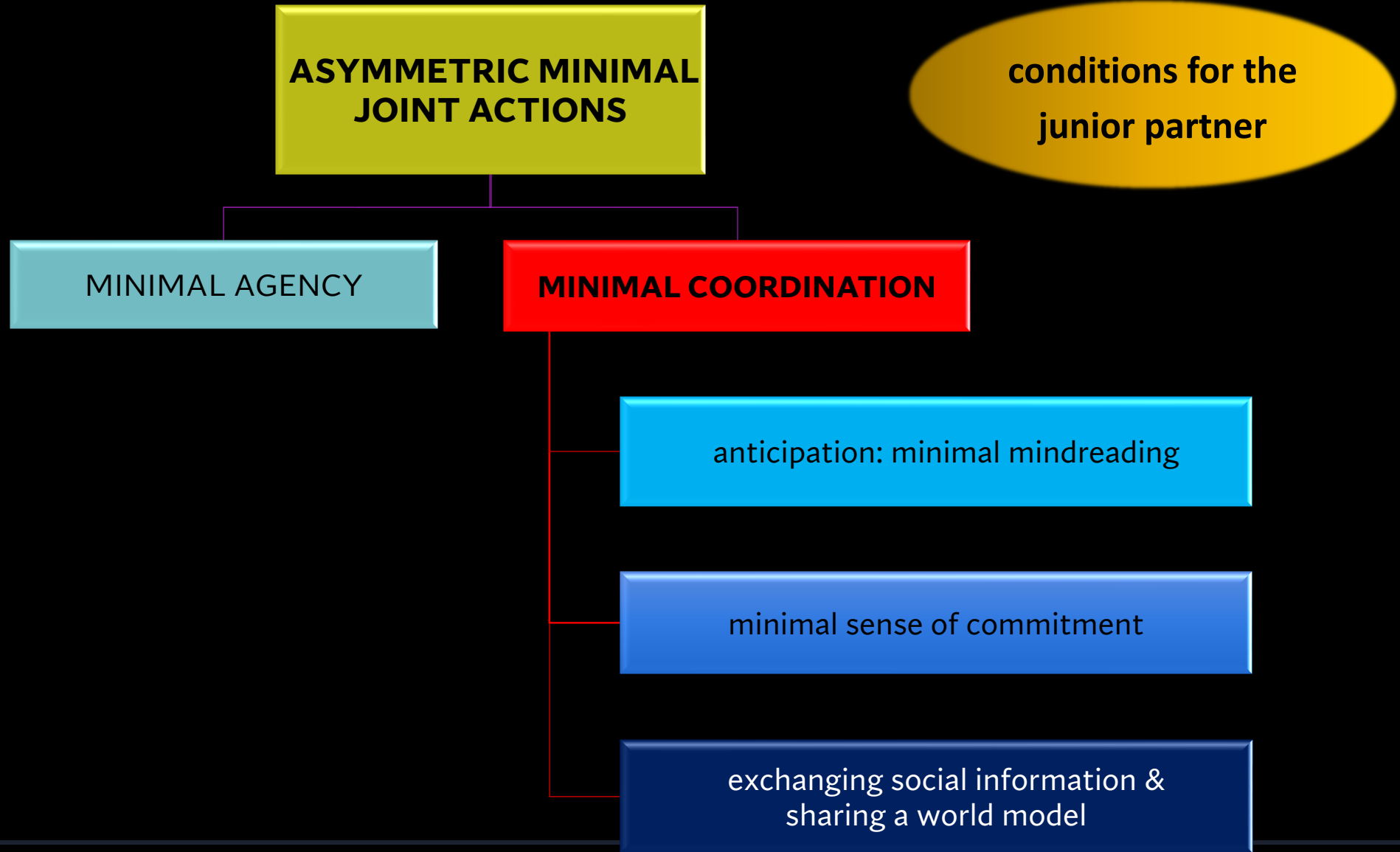
- have a difficulty explaining how one goes from one side of the divide to the other
- ❖ developmental & comparative psychology suggest that the change is *gradual and not sharp*

## The Biological Approach

- attempts to draw the distinction due to a mysterious capacity of our brain to generate consciousness, feeling, subjectivity, and meaning
- fail to explain what the missing quality is, how we can know when it is there and when it is missing
- ❖ why we should suppose that it can only be realized in **electro-chemical brain reactions**, and not in silicon systems, or **neural nets**

# Inbetween mere tool-use and social interactions

## MINIMAL COORDINATION



# Minimal mindreading



- ❖ **to coordinate your contribution in a joint action one has to be able to anticipate what the other agent will do next**



- utilize the notion of minimal mindreading that Steve Butterfill & Ian Apperley developed



- notion is a suitable starting point

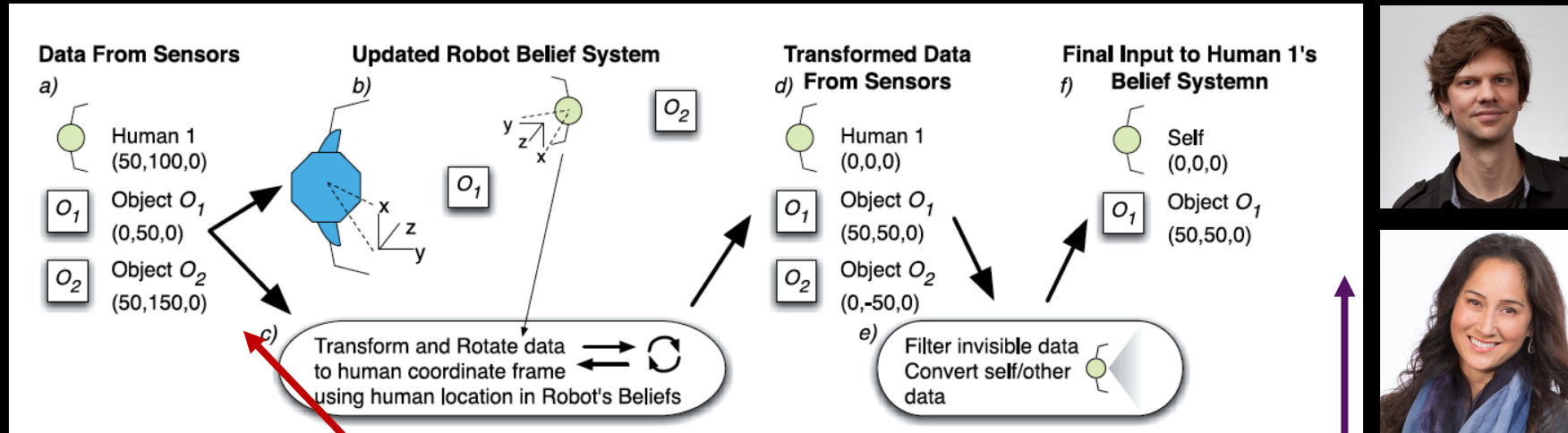
- as they claim that underlying processing are implicit, nonverbal, automatic, and based on unconscious reasoning



# Application to artificial agents

## MINIMAL MINDREADING

### MODELLING MENTAL STATES WITH RESPECT TO THE PERSPECTIVE OF THE HUMAN COUNTERPART



1. infer from their perception of the physical world to what a human counterpart can see or cannot see
2. infer that perspective of the human will guide future actions of the human

→ some cases of minimal mindreading can be achieved by artificial agents

# Minimal sense of commitment

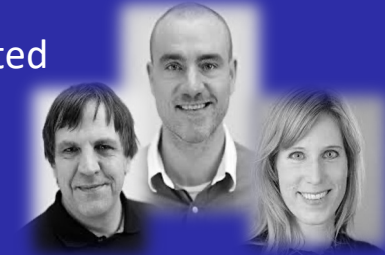


- ❖ 'social glue' for much of what counts as social interactions



- ❖ coordination abilities are also based on the capacity to form expectations and motivations with respect to your counterpart

- utilize the notion of a minimal sense of commitment that illuminates minimal forms of interpersonal commitments
  - components (expectation or motivation) of a standard commitment can be disassociated
    - single occurrence of just one component can be treated as a sufficient condition



- asymmetric joint actions:

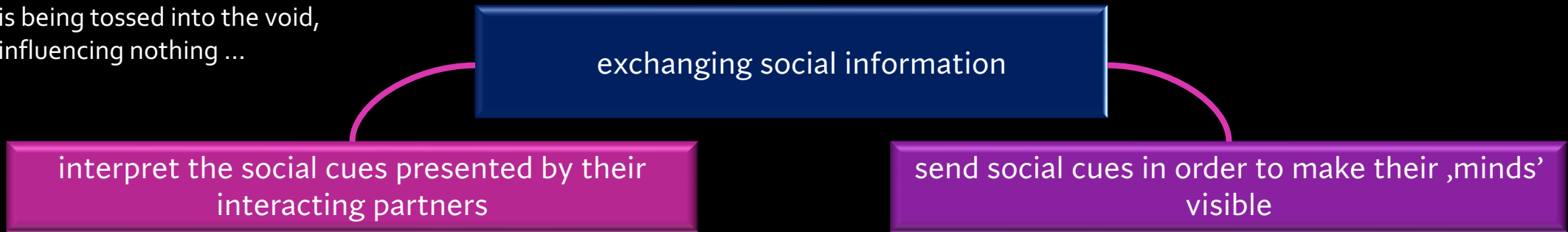
- minimal sense of commitment can be realized by just one participant
- most minimal case: only human counterparts entertain a minimal sense of commitment

# Exchanging social information

PROCESSING SOCIAL DATA PLAYS AN IMPORTANT ROLE IN SOCIAL INTERACTIONS

to exclude one-sided sociality, we need a minimal level of reciprocity

- social reaction to the Roomba is being tossed into the void, influencing nothing ...



**artificial agents are designed to enter the space of human social interaction**

- Wachsmuth: conversational agents like Max
- Kang et al.: gestures & emotional expression
- Baur et al.: ARIAS (artificial retrieval of information assistants) which are able to handle multimodal social interactions



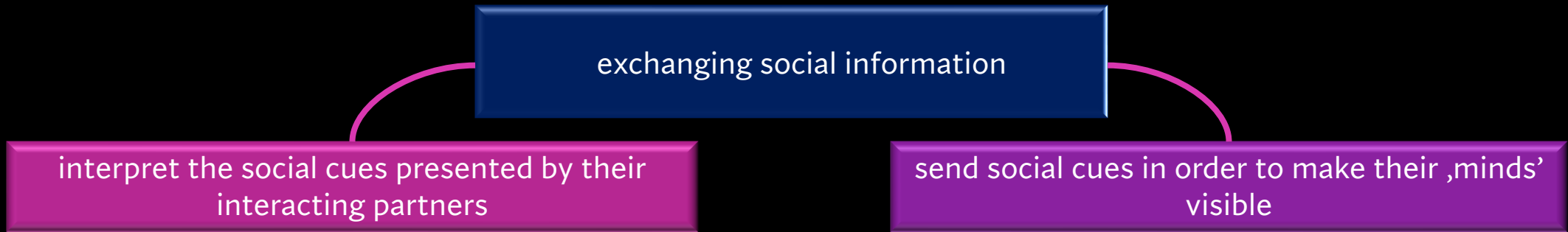
**→ react adequately to verbal and nonverbal behavior**

artificial agents are able to process social data & make use of it to anticipate the behavior of their interaction partners

# Exchanging social information

PROCESSING SOCIAL DATA PLAYS AN IMPORTANT ROLE IN SOCIAL INTERACTIONS

to exclude one-sided sociality, we need a minimal level of reciprocity



## LLMs

- apologies & social reactions are not being tossed into the void, they influence the machine's responses, and they do so in ways that make social sense
- Anger leads to apology. Questions lead to answers. Hints of sexual interest are picked up on and amplified back.

- You can productively take a social stance toward the machine
- You can call on your social skills in interacting with it
- you can coax the machine into further socially interpretable interactions

# Sharing a world

LLMs live NOT

**in our social, physical  
world**

**are not embodied**

But they may play a role in our  
world of language games.

## CONSIDER THE POSSIBILITY OF CHANGING THE WINNING TEAM & QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE ENTITIES

After all, we might be confronted with a new game.

- ❖ Before we can answer the question of **what we are doing when we interact with LLMs**, we have to conceptualize the INBETWEEN, because we cannot reduce our interactions with LLMs (and especially with future products of generative AI) to mere tool use.

**THE MAIN AIM OF THIS TALK WAS TO PREPARE THE GROUNDS FOR QUESTIONING THE DICHOTOMY BETWEEN ANIMATE AND INANIMATE ENTITIES, AS THIS IS AN IMPORTANT PRESUPPOSITION FOR ANY DEVELOPMENT OF NEW NOTIONS THAT CAN CAPTURE PHENOMENA THAT I LOCATE IN THE INBETWEEN.**

**IF WE ARE SUCCESSFUL WITH THIS, WE CAN ARGUE FOR A GRADUAL APPROACH.**

- machine designed in a way that exploits the fact that you will react to it as a social agent; and you, in turn, can exploit that fact about it.

# A gradual approach

## SINGLE-SIDED SOCIALITY

- sociality tossed into a void
- application of social skills
- reactions toward entities who are in no respect social partners, with no capacity for social uptake

## QUASI-SOCIALITY

- machines designed in a way that exploits the fact that you will react to it as a social agent; and you, in turn, can exploit that fact about it

## FULL-BLOWN, INTELLECTUALLY DEMANDING, COOPERATIVE SOCIAL INTERACTION

- both partners make second-order mental state attributions and satisfy various other conditions are required for full-blown adult human cooperative action

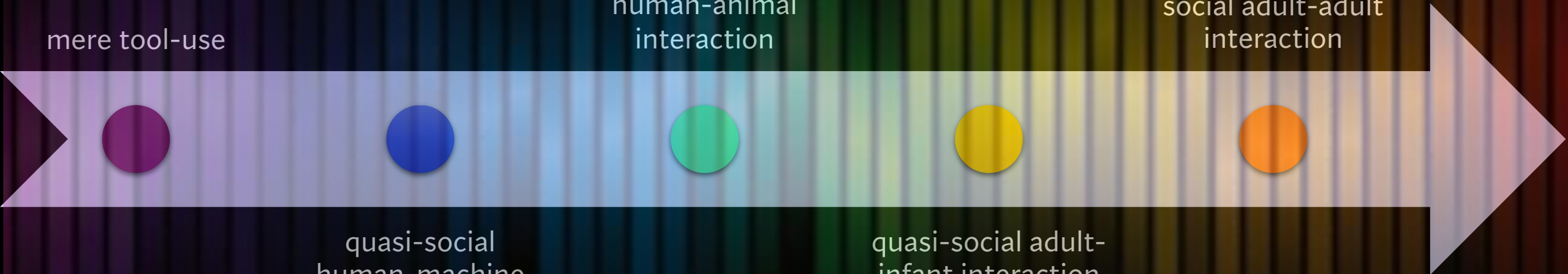
mere tool-use

quasi-social  
human-animal  
interaction

social adult-adult  
interaction

quasi-social  
human-machine  
interaction

quasi-social adult-  
infant interaction



All this would not have been possible if I had not interacted with people & machines



Daniel  
Dennett



Eric  
Schwitzgebel



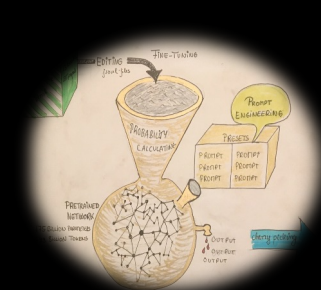
Mathew  
Crosby



David  
Schwitzgebel



Mike  
Wilby



DigiDan

Thank you!





# References

---

- Agrawal, A., Mackey, L., & Kalai, A. T. (2023). Do Language Models Know When They're Hallucinating References? (arXiv:2305.18248). *arXiv*. <http://arxiv.org/abs/2305.18248>
- Agüera y Arcas, B. (2022). Do Large Language Models Understand Us? *Daedalus*, 151(2), 183–197. [https://doi.org/10.1162/daed\\_a\\_01909](https://doi.org/10.1162/daed_a_01909)
- Barkham, P. (2021, July 25). Should rivers have the same rights as people? *The Guardian*. <https://www.theguardian.com/environment/2021/jul/25/rivers-around-the-world-rivers-are-gaining-the-same-legal-rights-as-people>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bunten, A., Iorns, C., Townsend, J., & Borrows, L. (2021, June 3). Rights for nature: How granting a river 'personhood' could help protect it. *The Conversation*. <http://theconversation.com/rights-for-nature-how-granting-a-river-personhood-could-help-protect-it-157117>
- Butterfill, S. A., & Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind & Language*, 28(5), 606–637. <https://doi.org/10.1111/mila.12036>
- Cole, S. (2023). 'It's Hurting Like Hell': AI Companion Users Are In Crisis, Reporting Sudden Sexual Rejection. *Vice*. <https://www.vice.com/en/article/y3py9j/ai-companion-replika-erotic-roleplay-updates>
- Dooley, B., & Ueno, H. (2022, April 24). This Man Married a Fictional Character. He'd Like You to Hear Him Out. *The New York Times*. <https://www.nytimes.com/2022/04/24/business/kihiko-kondo-fictional-character-relationships.html>
-

# References

---

- Gunkel, D. J. (2020). Robot Rights – Thinking the Unthinkable. In *Smart Technologies and Fundamental Rights* (pp. 48–72). Brill. [https://doi.org/10.1163/9789004437876\\_004](https://doi.org/10.1163/9789004437876_004)
- (2023). *Person, Thing, Robot: A Moral and Legal Ontology for the 21st Century and Beyond*. The MIT Press. <https://doi.org/10.7551/mitpress/14983.001.0001>
- Gray, J., & Breazeal, C. (2014). Manipulating Mental States Through Physical Action: A Self-as-Simulator Approach to Choosing Physical Actions Based on Mental State Outcomes. *International Journal of Social Robotics*, 6(3), 315–327. <https://doi.org/10.1007/s12369-014-0234-2>
- Henrich, J. P. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
- Heyes, C. (2014). False belief in infancy: A fresh look. *Developmental Science*, 17(5), 647–659. <https://doi.org/10.1111/desc.12148>
- (2015). Animal mindreading: What’s the problem? *Psychonomic Bulletin & Review*, 22(2), 313–327. <https://doi.org/10.3758/s13423-014-0704-4>
- Jensen, C. B., & Blok, A. (2013). Techno-animism in Japan: Shinto Cosmograms, Actor-network Theory, and the Enabling Powers of Non-human Agencies. *Theory, Culture & Society*, 30(2), 84–115. <https://doi.org/10.1177/0263276412456564>
- Kang, S.-H., Gratch, J., Sidner, C., Artstein, R., Huang, L., & Morency, L.-P. (2012). Towards building a Virtual Counselor: Modeling Nonverbal Behavior during Intimate Self-Disclosure. In Conitzer, Winikoff, Padgham, & van der Hoek (Eds.), *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*. [https://aamas.csc.liv.ac.uk/Proceedings/aamas2012/papers/2A\\_2.pdf](https://aamas.csc.liv.ac.uk/Proceedings/aamas2012/papers/2A_2.pdf)
-

# References

---

- Lake, B. M., & Baroni, M. (2023). Human-like Systematic Generalization through a Meta-learning Neural Network. *Nature*, 1–7. <https://doi.org/10.1038/s41586-023-06668-3>
- Lemoine, B. (2022, June 11). Is LaMDA Sentient? — An Interview. *Medium*. <https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917>
- Marcus, G., & Davis, E. (2020). GPT-3, Bloviator: OpenAI’s language generator has no idea what it’s talking about. *MIT Technology Review*. <https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion>
- Mattar, N., & Wachsmuth, I. (2012). Who are you? On the acquisition of information about people for an agent that remembers. *ICAART 2012 - Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, 2. <https://pub.uni-bielefeld.de/record/2440520>
- Michael, J., Sebanz, N., & Knoblich, G. (2016). The Sense of Commitment: A Minimal Approach. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01968>
- O’Donnell, E., & Talbot-Jones, J. (2017, March 23). Three rivers are now legally people – but that’s just the start of looking after them. *The Conversation*. <http://theconversation.com/three-rivers-are-now-legally-people-but-thats-just-the-start-of-looking-after-them-74983>
- Pacherie, E. (2013). Intentional joint agency: Shared intention lite. *Synthese*, 190(10), 1817–1839. <https://doi.org/10.1007/s11229-013-0263-7>
- Perner, J. (1991). *Understanding the representational mind*. The MIT Press.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>
-

# References

---

- Robertson, J. (2014). Human Rights vs. Robot Rights: Forecasts from Japan. *Critical Asian Studies*, 46(4), 571–598. <https://doi.org/10.1080/14672715.2014.960707>
- (2017). *Robo sapiens japonicus: Robots, Gender, Family, and the Japanese Nation*.
- Salmón, E. (2000). Kincentric Ecology: Indigenous Perceptions of the Human-Nature Relationship. *Ecological Applications*, 10(5), 1327–1332. <https://doi.org/10.2307/2641288>
- Scarborough, J. K., & Bailenson, J. N. (2014). Avatar Psychology. In M. Grimshaw (Ed.), *The Oxford Handbook of Virtuality*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199826162.013.033>
- Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. The MIT Press. <https://doi.org/10.7551/mitpress/9780262016797.001.0001>
- Strasser, A. (2006). *Kognition künstlicher Systeme*. De Gruyter. <https://doi.org/10.1515/9783110321104>
- (Ed.). (2024). *Anna's AI Anthology. How to live with smart machines?* xenomoi Verlag.
- Strasser, A., & Schwitzgebel, E. (2024). Quasi-sociality: Toward Asymmetric Joint Actions. In *Anna's AI Anthology. How to live with smart machines?* xenomoi Verlag.
- Strasser, A., & Wilby, M. (2023). The AI-Stance: Crossing the Terra Incognita of Human-Machine Interactions? In *Social Robots in Social Institutions* (pp. 286–295). IOS Press. <https://doi.org/10.3233/FAIA220628>
-

# References

---

- Tomasello, M. (2008). *Origins of human communication*. MIT Press.
- Trott, S., Jones, C., Chang, T., Michaelov, J., & Bergen, B. (2023). Do Large Language Models Know What Humans Know? *Cognitive Science*, 47(7), e13309. <https://doi.org/10.1111/cogs.13309>
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23(8), 998–1003. <https://doi.org/10.1016/j.neunet.2010.06.002>
- Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative Activities in Young Children and Chimpanzees. *Child Development*, 77(3), 640–663. <https://doi.org/10.1111/j.1467-8624.2006.00895.x>
- Weil, E. (2023, March 1). You Are Not a Parrot. *New York Magazine*. <https://nymag.com/intelligencer/article/ai-artificial-intelligence-chatbots-emily-m-bender.html>
- Wilby, M., & Strasser, A. (2024). Situating machines within normative practices: Bridging responsibility gaps with the AI-Stance. In A. Strasser (Ed.), *Anna's AI Anthology. How to live with smart machines?* xenomoi Verlag.
-