Dagstuhl Perspectives Workshop 24492

Human-in-the-Loop Learning through Grounded Interaction in Games

Raffaella Bernardi, Julia Hockenmaier, Udo Kruschwitz and Massimo Poesio

Motivation for this workshop

The convergence of three developments

- A decisive move in Artificial Intelligence (AI) towards human-centered intelligence and AI models that can learn through interaction and are able to act as intelligent assistants
- progress on grounded interaction – developing models aware of the situation in which they operate (a physical world in the case of robots, a virtual world in the case of artificial agents) and able to, e.g., understand / produce references to this situation
- The appearance of virtual world games in which conversational agents impersonating characters can learn to perform tasks, or improve their communicative ability, by interacting with human players in platforms such as Minecraft

Progress on intelligent assistants

> Hello. We are organising a Dagstuhl perspectives workshop on Human in the Loop Learning Through Grounded Interaction in Games, and we need some help.

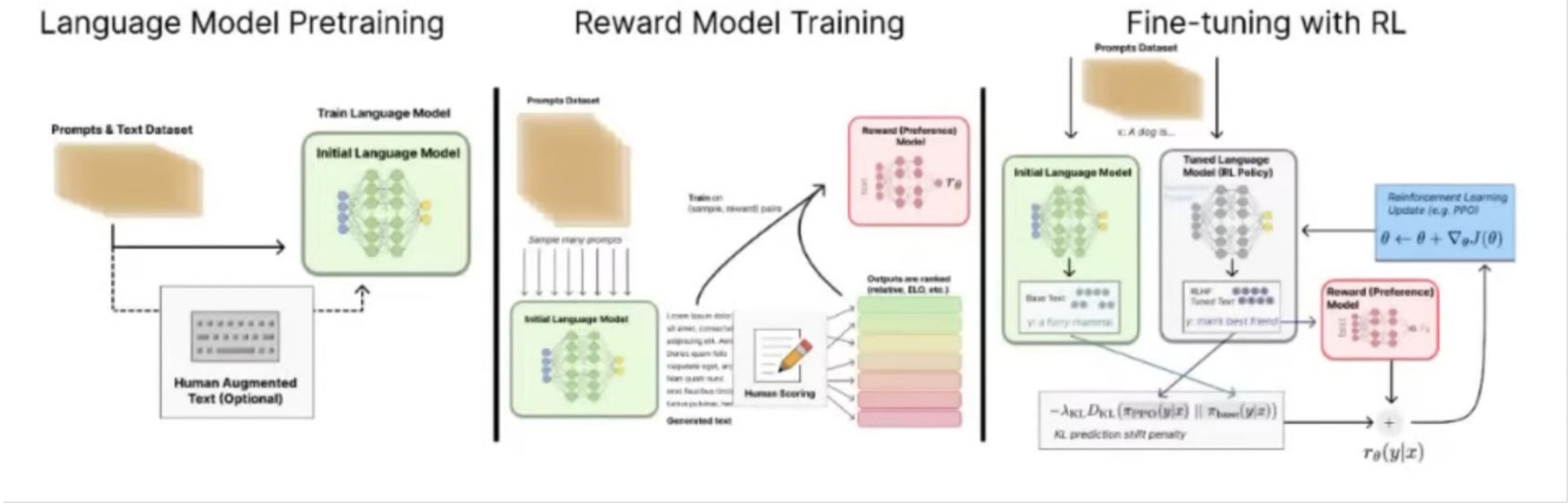
That sounds like an exciting workshop! I'd be happy to help. Could you share more details about what you need assistance with? Are you looking for help with organizing the event, structuring the workshop, developing content, or something else related to the topic of "Human in the Loop Learning Through Grounded Interaction in Games"?

> Well, first of all we need to discuss the state of the art on conversational assistants. What do you know about that?

Conversational assistants, often powered by artificial intelligence (AI) and machine learning (ML), have made significant advances in recent years, becoming a core part of many applications, ranging from personal assistants (like Siri, Al[illegible], and Google Assistant) to more specialized tools for customer service, healthcare, and entertainment. They interact with users via text or voice,

Multimodal LMs

Learning from human preferences (interactively?): RLHF



https://huggingface.co/blog/rlhf

Learning through interaction (online learning, continual learning, interactive learning, …)

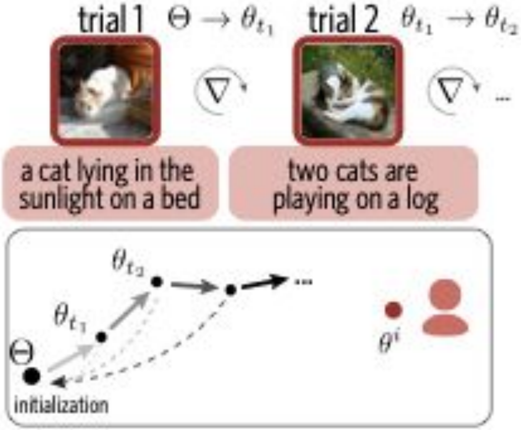


Figure 1: We introduce a regularized continual learning approach allowing agents initialized with a pre-trained language model Θ to iteratively infer the language model θ^i used by a partner, over repeated interactions $\{t_1, t_2 \dots\}$ in an online reference game.

Hawkins et al, 2020

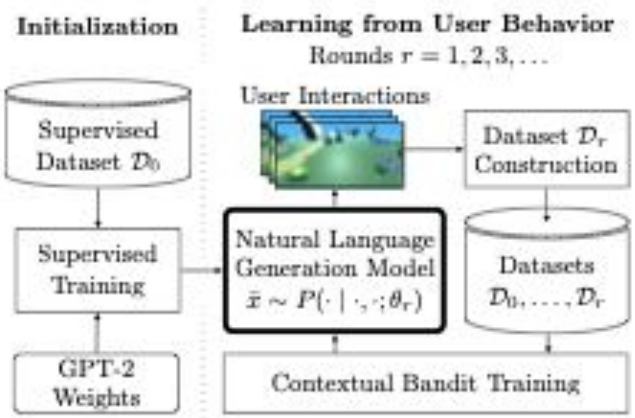


Figure 1: Diagram of our learning process. We initialize a generation model using supervised learning, and continually learn through interaction with users, by alternating between observing user execution of generated instructions and training.

Kojima, Suhr and Artzi, 2021

See also
Simpson et al 2020

Learning through interaction

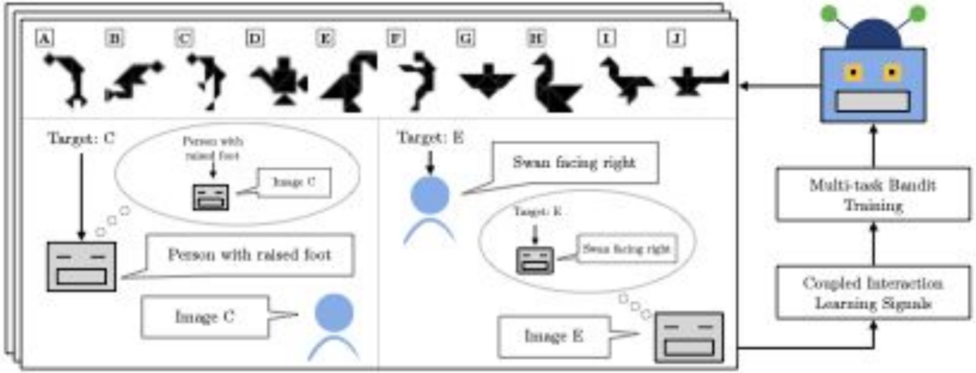


Figure 2: Illustration of our continual learning scenario with coupled comprehension and generation. The process alternates between interactions with human partners in a reference game, and training using learning signals from the interactions. The model performs both the generation (left) and comprehension (right) tasks, while jointly reasoning over the other role (thought bubbles). Training leverages feedback for the role the model performs as well as the opposing role. Following each round of training, we re-deploy the updated model and repeat the process.

(Gul and Artzi, 2024)

Grounded interaction

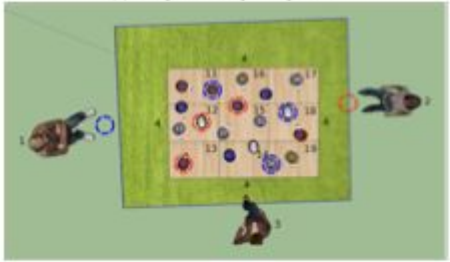
(a) GuessWhat?! scene

Figure 2: Game board in the RDG setting

(4) blue L on the top and the harry potter sign on the right

(a) Perspective of participant 1.

(b) Perspective of participant 2.

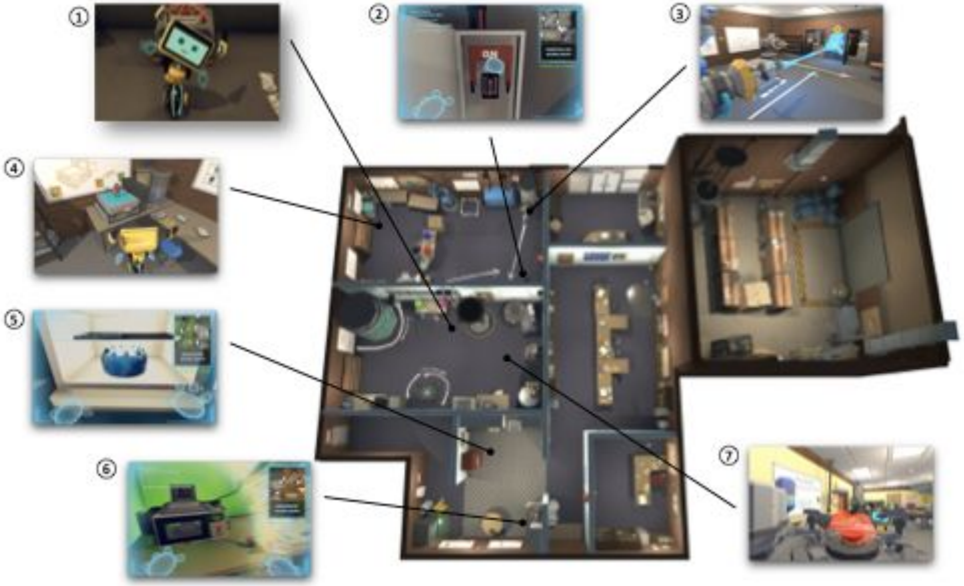


(c) Top-down perspective of the Cups corpus scene with ground truth object IDs.

Images from Suglia et al (2024) Zarriess et al (2016) Loaiciga et al (2021)

Grounded interaction

Much of this progress through multimodal environments



Images from Suglia et al (2024) Gao et al (2023)

Grounded interaction

And / or games

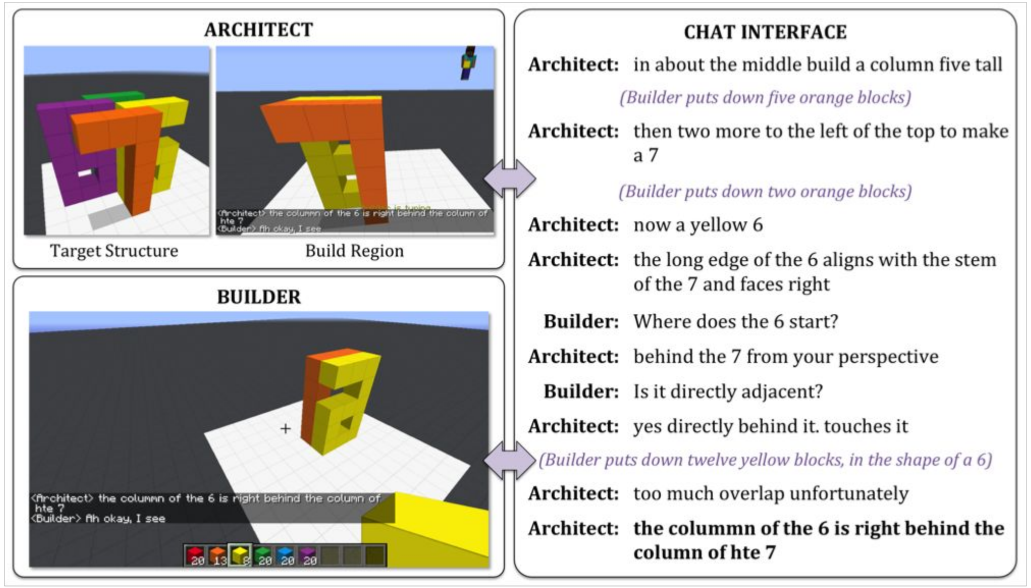


Figure 1: In the Minecraft Collaborative Building Task, the Architect (**A**) has to instruct a Builder (**B**) to build a target structure. **A** can observe **B**, but remains invisible to **B**. Both players communicate via a chat interface. (NB: We show **B**'s actions in the dialogue as a visual aid to the reader.)

Image from Narayan-Chen et al 2019

Open issues

Workshop motivations

- Discuss this progress in grounded interaction and learning from grounded interaction, take stock, and identify open questions

Where we are coming from

Julia:
Language and vision,
Embodied NLP
Minecraft

Massimo:
Reference, dialogue,
Human-in-the-loop in games
Multimodality in the brain
Most recently: ARCIDUCA

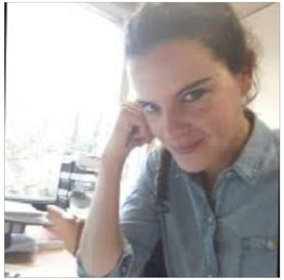
Raffaella:
Language and vision,
Visual question answering
Visual grounding,
Dialogue games

Udo:
Interactive IR
Human-in-the-loop in games
Downstream applications

The ARCIDUCA project

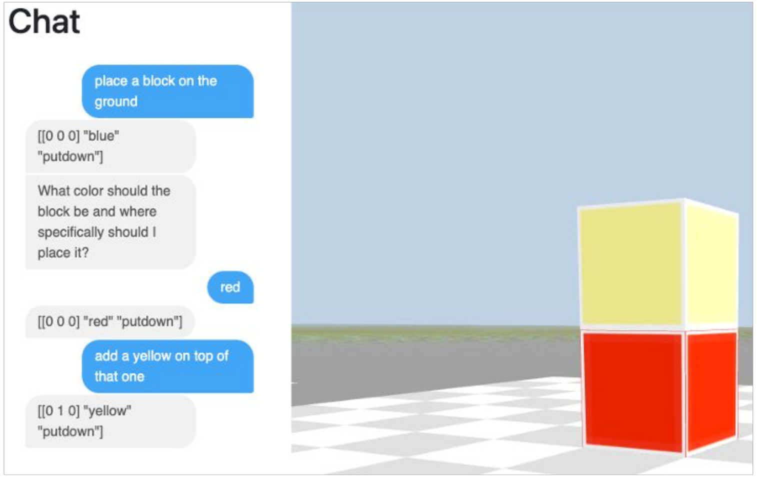
- A three-year project studying how conversational agents can learn through interaction in games
- Focusing in particular on reference
- Looking at two domains
  - Dungeons-and-dragons style textual games (Light)
  - Minecraft for 3D visual virtual world interaction




Conversational interaction in games

3D Visual: Minecraft



Allison et al 2017, 2018;
Narayan-Chen et al 2019;
Szlam et al 2019;
Kiseleva et al 2022; ….
Madge & Poesio 2024

Textual: Light, Dungeons and Dragons

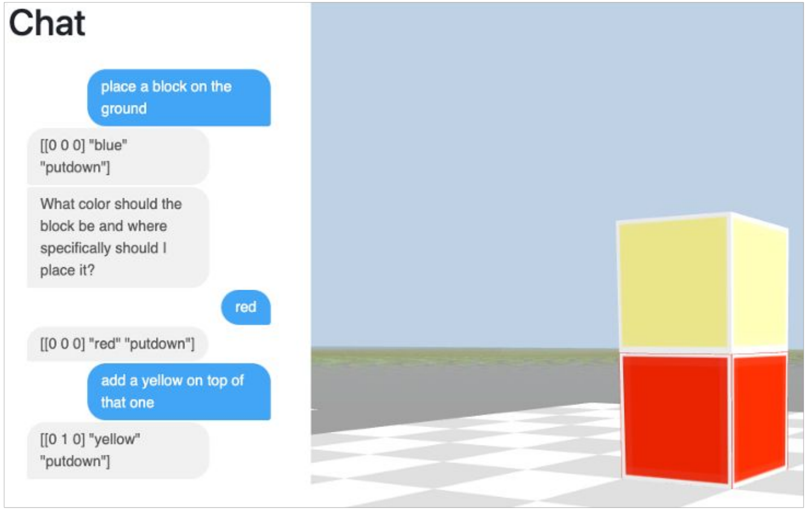
Persona 1: Servant	**Persona 2: King**
I come from the lower class. I do what I am told without question. I can not read. I have not seen my family in a long time.	I am a king of the whole empire. I give rules and pursuit them. I am brave and fearless.
Carrying: a duster, a small bucket, a rag Wearing: a shirt	Carrying: a crown, a scepter

Setting: The main foyer is massive. A grand staircase sits to the back of the foyer leading to the upstairs. At the front of the foyer stand two servants ready to help anyone who comes to visit. To the left of the room there is a doorway leading into a corridor. To the right there is a door leading to another corridor for the King's servants. At the foot of the stairs there is a bearskin rug that is staring at you almost as if still hungry. The walls are lined with portraits of the king and his family.

Servant: my humble king. What am I to do to serve you?
King: Ahhh. My loyal servant. Polish my scepter.
Action: give scepter to servant
Servant: Yes my lord. I will polish it immediately. Am I to return it to you personally?
Action: put scepter in small bucket
King: Yes. Yes. Of course. Also check the jewels in my crown. They seem loose.
Action: give crown to servant
Servant: But sire I am not qualified to do that. Would you prefer I take it to someone?
King: Oh fine then.
Emote: gesture sigh
Servant: I am sorry sir the rug startled me
Action: drop crown
King: Haha! That's bear I slain on my latest hunting trip. He's a mighty beast!
Emote: gesture laugh

Urbanek et al 2019,
Ammanabrolu et al 2021,
Shuster et al 2021, …

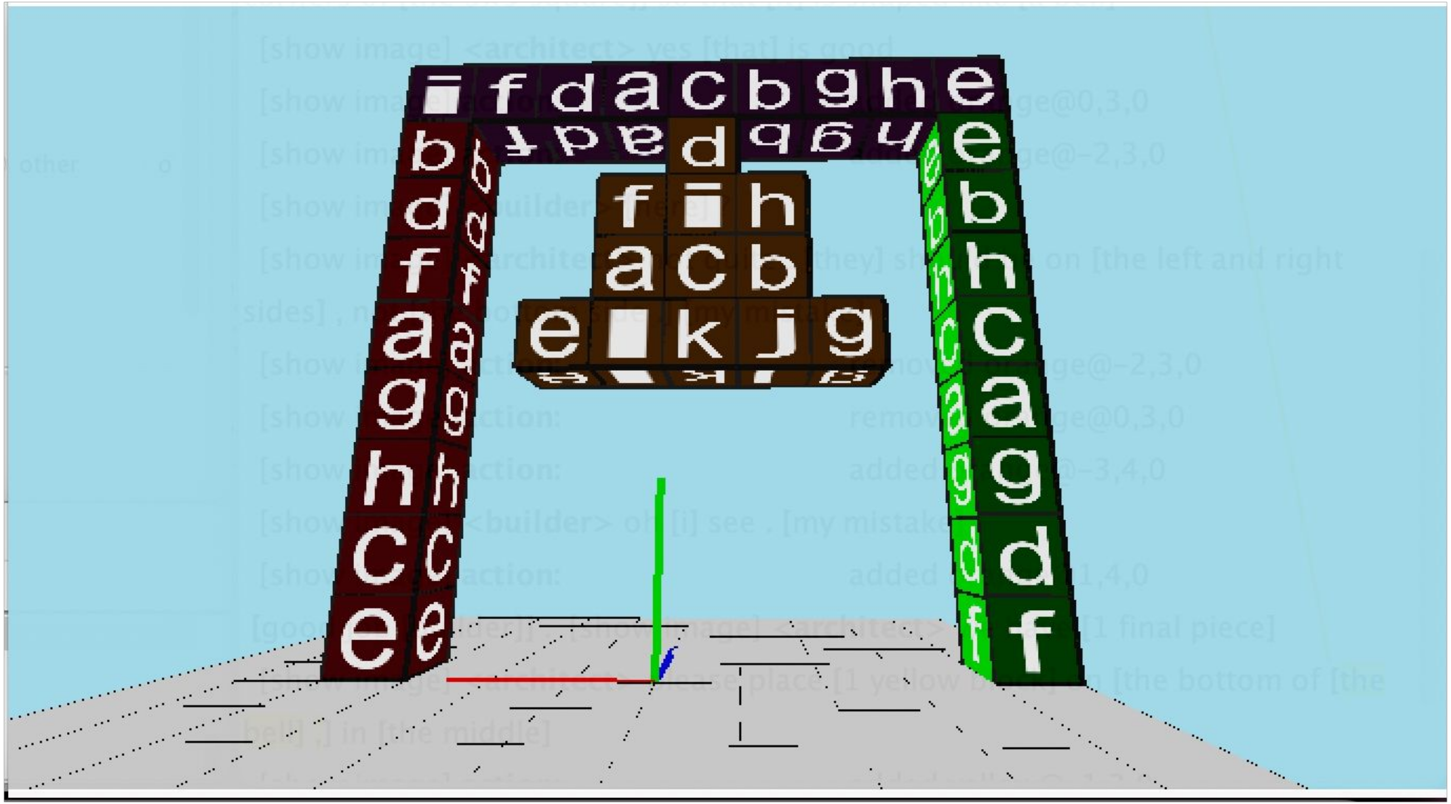
The LLMS Playing Minecraft platform



Madge & Poesio 2024

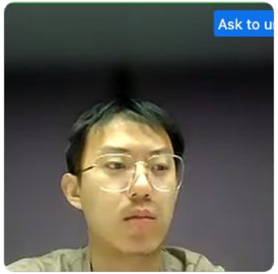
https://www.youtube.com/watch?v=N9n7u52Bbtk

The Minecraft Dialogue Corpus of Reference



Referring and grounding in games

- The CODI/CRAC 2022 corpus (Yu et al, CODI/CRAC 2022)
  - Five conversational datasets annotated with anaphoric and deictic reference according to an updated version of the ARRAU guidelines, including a part of the Light corpus
- Coreference with LLMs (Gan, Yu and Poesio, LREC 2024; Yang)
- Reference resolution in the MDC-R (Shao)
  - Conceptual pacts (Hough et al, 2024)
- Benchmarks for testing LLM's ability to ask clarifications (Gan, Purver and Poesio, SIGDIAL 2024; Gan, Purver and Poesio, under review)



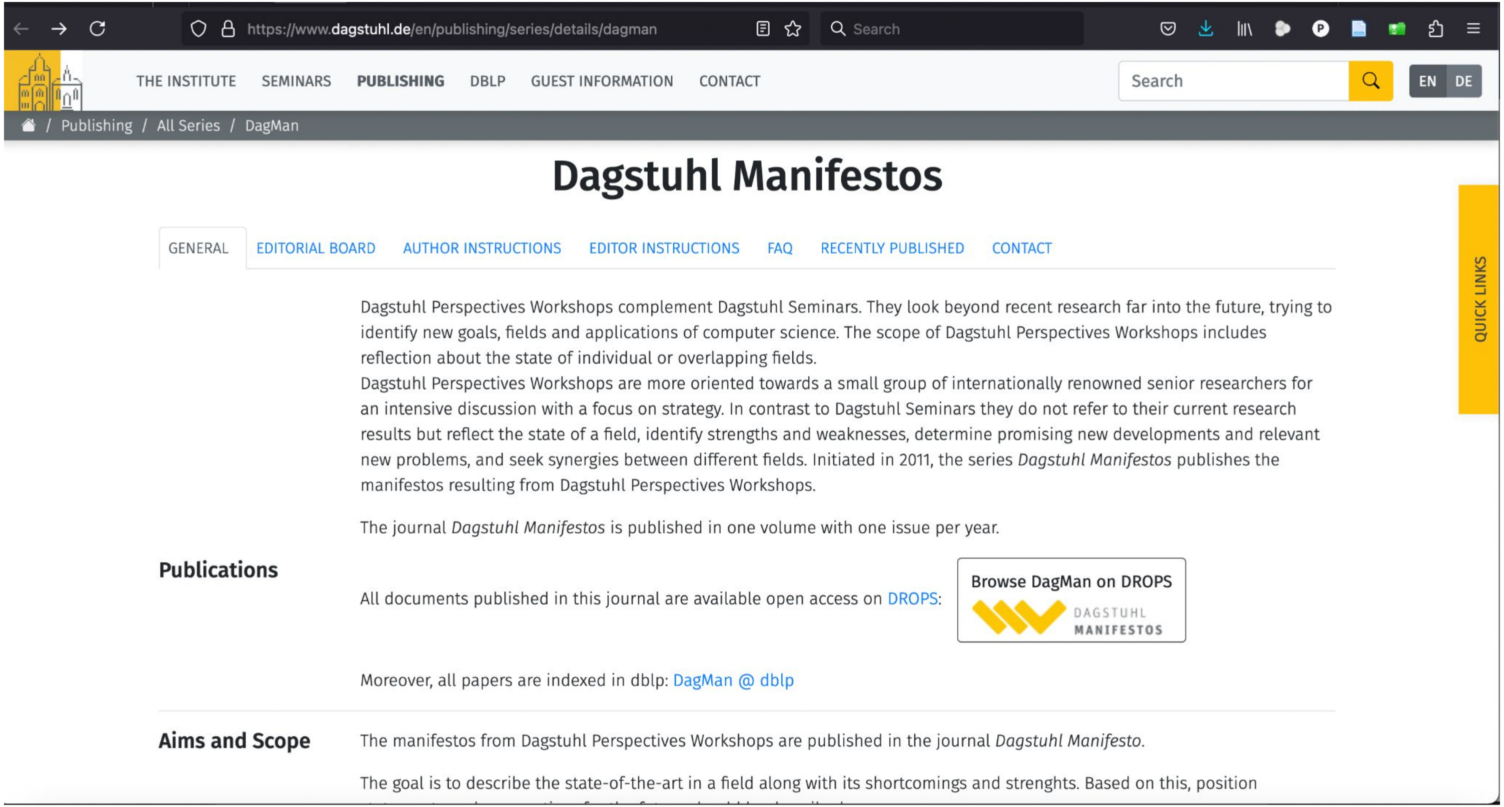
Research questions for the workshop

Can we improve the grounded communication and task performance abilities of (embodied/multimodal, conversational) AI agents through the use of games and human-in-the-loop learning?

Objectives for the week

- Like every other Dagstuhl workshop, we should produce a **report** **https://www.dagstuhl.de/en/publishing/series/details/DagRep**
- The **distinctive output** of a Perspectives workshop should be a **Dagstuhl Manifesto** ("in addition to the report") that "should include research directions that are put into a larger context, like its relevance for society and economy, applications, and relations to other fields. Its audience goes beyond the inner circle of experts and should include policy makers." https://www.dagstuhl.de/en/publishing/series/details/dagman
- An unofficial objective of the workshop will be to (start designing) **new tasks** that could inspire the new research directions identified in the manifesto, and perhaps identify possible **collaborations**.

The Dagstuhl Manifestos



https://www.dagstuhl.de/en/publishing/series/details/dagman

THE INSTITUTE SEMINARS PUBLISHING DBLP GUEST INFORMATION CONTACT

Search EN DE

Publishing / All Series / DagMan

QUICK LINKS

Dagstuhl Manifestos

GENERAL EDITORIAL BOARD AUTHOR INSTRUCTIONS EDITOR INSTRUCTIONS FAQ RECENTLY PUBLISHED CONTACT

Dagstuhl Perspectives Workshops complement Dagstuhl Seminars. They look beyond recent research far into the future, trying to identify new goals, fields and applications of computer science. The scope of Dagstuhl Perspectives Workshops includes reflection about the state of individual or overlapping fields.
Dagstuhl Perspectives Workshops are more oriented towards a small group of internationally renowned senior researchers for an intensive discussion with a focus on strategy. In contrast to Dagstuhl Seminars they do not refer to their current research results but reflect the state of a field, identify strengths and weaknesses, determine promising new developments and relevant new problems, and seek synergies between different fields. Initiated in 2011, the series *Dagstuhl Manifestos* publishes the manifestos resulting from Dagstuhl Perspectives Workshops.

The journal *Dagstuhl Manifestos* is published in one volume with one issue per year.

Publications

All documents published in this journal are available open access on DROPS:

Browse DagMan on DROPS
DAGSTUHL MANIFESTOS

Moreover, all papers are indexed in dblp: DagMan @ dblp

Aims and Scope

The manifestos from Dagstuhl Perspectives Workshops are published in the journal *Dagstuhl Manifesto*.

The goal is to describe the state-of-the-art in a field along with its shortcomings and strenghts. Based on this, position

Schedule

Monday identifying **strengths and shortcomings** of current state of the art

09:00-10:30 Intro to the seminar (30 min), participants' short self-introduction (1 hr)
10:30:11:00 coffee
11:00-12:00 Presentation SOTA 1: What are the grounded and communication task performance abilities of (embodied/multimodal, conversational=emc) AI agents
LUNCH
14:00-15:00 Presentation SOTA 2: What kind of games / multimodal platforms are currently used to evaluate / train e/m/c AI agents?
15:00-16:00 Coffee Break
16:00-17:00 Presentation SOTA 3: What are the current approaches to (human vs. artificial agents in the loop) learning for AI agents?
17:00-18:00 General Discussion

Schedule

Tuesday Identifying **development perspectives** of Grounded Interaction

09:00-10:00 **Summary of SOTA** (Massimo, Udo, Julia, Raffa)

10:30-12:00 5 Presentations on existing efforts to go beyond the SOTA

- Bernardi, Fernandez, Koller, Schlangen, Suglia. The PlayPen shared task.
- Malihe Alikhani. TBC
- Marc-Alexandre Cote. TW-Bench: Text-games Benchmark.
- Albert Gatt and Sina Zarriess. Reference Games.
- Julia Hockenmaier, Prashant Jayannavar, Chris Madge, Massimo Poesio, Nicholas Asher, … . Minecraft

14:00-15:00 General Discussion

15:00-16:00 Coffee Break

16:00-17:00 Collaboration time [we prepare the summary]

17:00-17:30 **Summary of development perspective** (Massimo, Udo, Julia, Raffa)

17:30-18:00 Creation of working groups

Schedule

Wednesday **Directions for New Research Projects/New Tasks**
09:00-10:00 Definition of working groups
10:00-12:00 Breakup into WG
PM Trip to Trier Weihnachtmarkt

Thursday **Directions for New Research Projects/New Tasks**
09:00-12:00 WG
14:00-15:00 Report of the Working Group/Feedback
15:00-16:00 Coffee Break
16:00-17:00 Summary **on new tasks** (Massimo, Udo, Julia, Raffa)
17:00-18:00 General Discussion

Friday **Manifesto**
AM Working Group on the Manifesto section
12:00 END OF THE MEETING
PM Massimo, Udo, Julia, Raffa work on the Manifesto

SOTA Presentations

On Monday, we had three SOTA Presentations

- SOTA 1: What are the grounded and communication task performance abilities of (embodied/multimodal, conversational) AI agents (Julia Hockenmaier, Raquel Fernandez, Massimo Poesio, Raffaella Bernardi, Sina Zarrieß…)
- SOTA 2: What kind of games/multimodal platforms are currently used to evaluate/train embodied/multimodal, conversational AI agents? (R. Bernardi, M.-A. Côté, D. Perez-Liebana, D. Schlangen and A.Suglia)
- SOTA 3: Current Approaches to Human (and Artificial Agent?)-in-the-Loop Learning for AI Agents (A. Suglia, M. Alikhani, M.-A. Côté, A. Koller, E. Simpson, D. Schlangen, A. Suhr)

On Tuesday, we had several more presentations on individual projects

- PlayPen, Gatt & Zarrieß, Koller, Illinikh & Loaiciga, Suhr, Cote on TW-Bench, Alikhani, Hockenmaier et al on the Minecraft Dialogue Corpus

Promising Research Directions

We identified four research directions, and the participants were divided in four WGs discussing what kind of research should be carried out in these:

- WG1: Complex Interaction (the ‘Karlsruhe’ group)
- WG2: Perceptual grounding and Embodiment (the ‘S003’ group)
- WG3: Game design for interaction modelling (the ‘Wine Cellar’ group)
- WG4: Learning from Interaction (the ‘Kaiserslautern’ group)

WG1 : Complex Interaction. Motivations

A common assumption in many computational models is that dialogue consists of a linear sequence of turns in which two agents alternately exchange information. Each turn is assumed to depend only on the last turn of the other participant. However, human conversation requires more complex forms of interaction spanning multiple turns to solve real-world tasks.

Complexity occurs for several reasons: Dialogue is done by multiple people. They start from different information states, have different perspectives, and cannot see what is in each other's minds. They have a social relationship that they have to manage. Their interaction happens in real time, across multiple modalities, in the presence not only of various kinds of noise but of fundamental asymmetries in what the participants can perceive, know and understand.

To successfully overcome these asymmetries and solve tasks through such interactions, the interaction scheme needs to offer a number of functions (see below). Among humans, these are exemplified by a variety of phenomena that depend not just on sequential information exchange but on more complex structures, with richer models of the local and global interaction context. It is not clear to what degree current LLM-based models of dialogue can cope with them, and how much this limits their ability to collaborate efficiently with humans.

WG1: Complex Interaction. Questions and outcome

Questions for the WG

- How can we train models to participate in complex interactions?
- What can we learn from complex interactions?

Outcome:

- A fairly detailed list of ‘complex interaction’ issues (fed into the ‘Learning from Interaction’ WG)

WG2: Perceptual Grounding and Embodiment

Premise: Compare an agent’s “mental model” of a situation when having different roles within a game (assess its “understanding (grounding)”)

1. An observer of a multi-agent situation (creating mental model via perception (and language?))
2. An actor working in isolation (manipulating the environment, i.e., enriching the mental model via embodiment/embodied experience)
3. A game master of a multi-agent situation
4. An actor working within a multi-agent situation

Mental model: internal representation/understanding of a situation in an environment and of the involved objects and participants, their roles and interaction with each other

WG2: Perceptual Grounding and Embodiment

Questions for the working group:

- How beneficial is learning from grounded interaction?
- Grounding capabilities (NB related to systems’ ability to recognize their/others’ uncertainty)

Outcomes:

- The types of common ground required for different types of agent situations
- Different types of evaluation

WG3: Research direction(s) investigated

- Q1: How can games and game benchmarks be designed more systematically, such that they lead to a deeper understanding of games and the skills that games are testing? How do we generalize skills and agents' abilities across games?
- Q2: How important it is to maximize human engagement and fun in the design of these games?
- Q3: What role does the complexity of the game have? And how do we measure it?
- Q4: How do we evaluate agents within and across games? In particular, how do we evaluate whether the skills trained / tested with a game transfer to real world applications?

WG3: Outcomes

- A Taxonomy of skills
- The importance of complexity
- Fun / engagement
- Evaluation

WG3 Outcomes: Skills in games

	A	B	E	F	G	H	I	J	K	L	M
1		**Games:**	Love Letter	Catan	Pictionary	Daybreak	Mysterium	Gloom	We didn't playterst this at all	Pandemic	The Mind
2	Match	**Features / Skills**	4	4	4	4	5	5	6	6	6
3	x	Fully Cooperative				x	x			x	x
4	x	Teams (Coop + Comp)									
5		Dealing with Stochasticity	x	x		x		x	x	x	x
6		Dealing with Partial Observability	x	x		x		x	x		x
7		Understanding of Counting, Probabilities								x	x
8		Planning		x						x	x
9		Discussion (to determine action)				x	x			x	
10		Discussion (to identify knowledge)			x						
11		Voting (to perform actions)									
12		Theory of Mind			x			x			x
13		Language and Vision					x		x		
14		Negotiation (trading, agreements)		x							
15		Hidden identities									

WG3 Outcomes: More complex games

- More complex games allow for testing multiple skills at once. Example:
    - “Mascarade”: theory of mind, information gathering, deception, memory
    - “Codenames”: theory of mind, discussion, limited verbal action space.
    - “Gloom”: theory of mind, narrative, memory
    - Taxonomies that maps N games to M skills would help determine which games to use to evaluate a skill or sets of skills.
    - Allows for a degree of complexity.

- Using existing (real) games are more engaging environments and closer to real-world (see below).

WG3 Outcomes: Complex vs. simple games

- Q1: Do you need more complexity for proper testing?
- Q2: Is there a link between complexity and transferability?
  - E.g., are real world games better for training an agent to perform a given task?
- Note: there are different dimensions to complexity
  - real world vs puzzles
  - Even within puzzles / real world games you can you have degrees of complexity
    - E.g., in a Minecraft game you may have to do simple actions no interaction / you can start interacting with other people / the interaction may become arbitrarily complex
    - Cfr MeetUp
  - Do you need different games for testing complexity or just different degrees of complexity in single game?
- Issue: Full world games are more difficult to develop?
  - But: LLMs and platforms like Genie2 are making this easier
- Recommendations:
  - More research on Q1 and Q2
  - Be more explicit / aware about complexity in your game and how that may affect its usefulness

WG3: Fun/Engagement

- Gap: are current games for testing / training agents in HIL contexts fun?
  - Check current benchmarks
- Objective: develop fun games for testing / training agents
  - Reasons: this makes recruitment easier / collected data has better quality (also because players are more distracted from real task)
  - Particularly important in human in the loop situations, less or not at all in agent in the loop?
- Issue: How do you develop a game that is fun but also tests important skills
  - A big issue in literature on games with a purpose / serious games
- Ways to make a game engaging:
  - Game player design: be aware of player types (e.g., Bartle taxonomy) and try to target a specific player interest / motivation
  - Agent have personas with intentions goals
  - Achieving flow (Connection with game complexity)
  - Real-world physics
- Recommendations:
  - Community should test whether there is a connection between engagement and effectiveness of a benchmark re: testing / training
  - Should aim to develop engaging games (e.g., collaborating with game designers)

WG3: Fun / engagement 2

- Are current benchmarks fun?
    - Dataset-based benchmarks (e.g., HOLMES, Clembench): not applicable
    - Also not relevant for agent-in-the-loop learning benchmarks like KinderLLM
    - Are there any HIL benchmarks for testing CAs? Are any of these game-based?
        - TW-World / DiscoveryWorld /
        - ClarQ-LLM (Gan et al 2024)
        - PORTAL (possible to participate?)
        - HumBEL (Sicilia et al 2024) (NB: not really a game)

WG4

WG4: Questions/topics

- Continual learning vs. multi-turn adaptation/entrainment/working memory
- What kind of learning architecture is more appropriate for learning from interaction?
- Learning from scratch through interaction
- Learn through immediate feedback
- Can they collect information through conversation
- Ask informative questions, multi-turn, uncertainty
- Adaptive Learning
- Curriculum learning vs Continual learning
- Human in the loop vs agent in the loop
- Reasoning in interaction

What can we gain from learning from interaction?

- Discovering nuanced user goals, linguistic styles, and preferences
- Identifying systematic misunderstandings to improve factual accuracy and coherence
- Adapting to domain specific requirements within multi-turn interactions
- Resolve uncertainty

Dimensions of Language Learning

Dimension	Human Development	Machine Learning	Type of Data
Learning to have / to recognise intentions	Infant		Pointing and identification game
Learning to communicate	Infant		Multi-player games; actions can communicate
Learning a first language	Toddler, Child	Pre-training	General corpora, caregiver linguistic material
Learning to use language	//	Post-training: Instruction tuning, RLHF	Instruction prompts + replies; preference information
Adapting to partner	Child; ~ 4 yrs (acquisition); active in any conversation	//	
Learning non-language facts via language	Child; 4 yrs – 99 yrs	Post-training: Model editing	

Levels of "interactiveness"

Method		**Examples/Citations**
Autoregressive / behavior cloning / next token prediction		
Reinforcement Learning from Human Feedback (RLHF)	Generate version A, generate version B, let a preference model (learned from humans) rank them	**InstructGPT**
Direct preference optimisation (DPO)	Generate version A, generate version B, push up the (human) preferred one, push down the other	**Mazzacara et al. 2024**
Kahneman-Tversky optimisation (KTO)	Generate a continuation, humans provide thumb up/down Kahneman-Tversky model of human value, which allows us to directly optimize for utility instead of maximizing the log-likelihood of preferences	**Chen et al. 2024**
Monte-carlo Tree-Search	Explore different options and select (via some value function) the one that is more desirable long term	Agent Q
In-context Learning / Iterative refinement ("agent memory")		Brown et al. 2020 Park et al. 2023

Is learning to formulate intentions required for learning to communicate?

- Should agent have intentions?
  - Their own? Imposed upon them by a human who delegates to them tasks?
- Are LLM even able to formulate them? They are simply text generation machines!
- How about recognizing them? Is intent recognition required for task completion in agent settings?
- Is there a relationship with Theory-of-Mind? Again probably it doesn't make sense to talk about theory of mind for LLMs at all.
- Is there a relationship with common ground?

Overall, seems this is something that cannot be really represented in current models.

What are the language games that would facilitate this learning?

- This is very related with the working group designing games
- Transfer of models between games and environments
- Requirements on datasets, dataset creation, feature representations (from real world to the ML methods)
    - What can be sampled and how is it structured for different methods?

Next steps

- Official outcomes
  - We just produced a report summarizing what happened and the work of the WGs
  - We are starting
- Other outcomes
  - The PlayPen group just released a paper on their latest taxonomy of skills that an agent must have in order to interact successfully and a benchmark testing these skills
    - https://arxiv.org/abs/2502.14359
  - Some of the ideas / collaborations are feeding into the Gravitation proposal
    - E.g., Oertel
  - Other funding ideas