

Wordplay 2022

The 3rd Wordplay: When Language Meets Games Workshop

Proceedings of the Workshop

July 14, 2022

©2022 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-955917-81-0

Introduction

The Wordplay workshop focuses on exploring the utility of interactive narratives, think everything from classic text-adventures like Zork to modern Twine games, to fill a role as the learning environments of choice for language-based tasks including but not limited to storytelling. A few previous iterations of this workshop took place very successfully with hundreds of attendees, at NeurIPS 2018 and NeurIPS 2020. Since then, the community of people working in this area has rapidly increased. This workshop aims to be a centralized place where all researchers involved across a breadth of fields can interact and learn from each other. Furthermore, it acts as a showcase to the wider NLP/RL/Game communities on interactive narrative's place as a learning environment. The program features a collection of invited talks in addition to contributed posters from each of these sections of the interactive narrative community and the wider NLP and RL communities.

For more information visit: <https://wordplay-workshop.github.io/>.

-The Wordplay's Organizing Committee

Organizing Committee

Organizing Committee

Prithviraj Ammanabrolu, Allen Institute for AI
Xingdi Yuan, Microsoft Research
Marc-Alexandre Côté, Microsoft Research
Ashutosh Adhikari, Microsoft Turing

Advisory Committee

Michiaki Tatsubori, IBM Research
Matthew Hausknecht, Argo AI
Kory Mathewson, DeepMind
Jack Urbanek, Meta AI Research
Adam Trischler, Microsoft Research
Jason Weston, Meta AI Research

Program Committee

Additional Reviewers

Guiliang Liu, University of Waterloo & Vector Institute
Raghuram Mandyam Annasamy, Google

Invited Speakers

Shrimai Prabhumoye, Nvidia
Tim Rocktäschel, University College London & Deepmind
Hal Daumé III, University of Maryland & Microsoft Research
Peter Jansen, University of Arizona
Lynn Cherny, GhostWeather

Keynote Talk: Controllable Text Generation for Interactive Virtual Environments

Shrimai Prabhume
Nvidia

Abstract: Since the dawn of the digital age, interactive virtual environments and electronic games have played a huge role in shaping our lives. Not only are they a source of entertainment but they also teach us important life skills such as strategic planning, collaboration, and problem solving. Therefore, online gamers expect their virtual environment to be aware of their situation (e.g., position in a game) and interact with them in natural language. In this talk, I describe novel techniques to generate text in a particular style. This talk provides an approach of generating engaging naturalistic conversation responses using knowledge generated by pre-trained language models, considering their recent success in a multitude of NLP tasks. The talk will conclude with exploring whether pretrained language models can be situated in these virtual spaces and generate dialogue in a zero-shot manner.

Bio: Shrimai Prabhume is a research scientist at Nvidia. She got her PhD degree in computer science from the Language Technologies Institute, Carnegie Mellon University. She was advised by Prof. Alan Black and Prof. Ruslan Salakhutdinov. She work on controllable text generation with focus on style, content and structure. She is also exploring the ethical considerations of controllable text generation. She co-designed the Computational Ethics for NLP course which was offered for the first time in Spring 2018 at CMU.

Keynote Talk: Knowledge Intensive Reinforcement Learning

Tim Rocktäschel

University College London & Deepmind

Abstract: Progress in Reinforcement Learning (RL) methods goes hand-in-hand with the development of challenging environments that test the limits of current approaches. While existing RL environments are either sufficiently complex or based on fast simulation, they are rarely both these things. Moreover, research in RL has predominantly focused on environments that can be approached *tabula rasa*, i.e., without agents requiring transfer of any domain or world knowledge outside of the simulated environment. I will talk about the NetHack Learning Environment (NLE), a scalable, procedurally generated, stochastic, rich, and challenging environment for research based on the popular single-player terminal-based rogue-like game, NetHack. We argue that NetHack is sufficiently complex to drive long-term research on problems such as exploration, planning, skill acquisition, and language-conditioned RL, while dramatically reducing the computational resources required to gather a large amount of experience. Interestingly, this game is extremely challenging even for human players who often need many years to solve it the first time and who generally consult external natural language knowledge sources like the NetHack Wiki to improve their skills. I will cover some of our recent work on utilizing language information in this challenging environment.

Bio: Tim Rocktäschel is a Researcher at DeepMind and an Associate Professor at the Centre for Artificial Intelligence at University College London (UCL). There, Tim is leading the UCL Deciding, Acting, and Reasoning with Knowledge (DARK) Lab. Tim is also a Scholar of the European Laboratory for Learning and Intelligent Systems (ELLIS). Before, he was a Manager, Research Scientist, and Area Lead at Meta AI (FAIR), a Postdoctoral Researcher in Reinforcement Learning at the Whiteson Research Lab at the University of Oxford, a Junior Research Fellow in Computer Science at Jesus College, and a Stipendiary Lecturer in Computer Science at Hertford College. Tim obtained his Ph.D. from UCL under the supervision of Sebastian Riedel, where he was awarded a Microsoft Research Ph.D. Scholarship in 2013 and a Google Ph.D. Fellowship in 2017. Tim's work focuses on Reinforcement Learning and Open-endedness.

Keynote Talk: Training Agents to Learn to Ask for Help in Virtual Environments

Hal Daumé III

University of Maryland & Microsoft Research

Abstract: Agent has largely become synonymous with autonomous agent, but I'll argue that scoping our study of agents to those that are fully autonomous is a mistake: instead, we should aim to train agents that can assist humans, and be assisted by humans. In line with this goal, I will describe recent and ongoing work in the space of assisted agent navigation, where agents can ask humans for help, and where they can describe their own behaviors. This talk will largely be based on joint work with Khanh Nguyen and Lingjun Zhao.

Bio: Hal Daumé III is a professor at University of Maryland, in the computer science department. He is also a Senior Principal Researcher at Microsoft Research NYC. Hal's main focuses are in natural language processing and machine learning; he studies questions related to how to get machines to become more adept at human language, by developing models and algorithms that allow them to learn from data. The two major questions that really drive their research these days are: (1) how can we get computers to learn through natural interaction with people/users? and (2) how can we do this in a way that minimize harms in the learned models?

Keynote Talk: ScienceWorld: Is your Agent Smarter than a 5th Grader?

Peter Jansen
University of Arizona

Abstract: Question answering models have rapidly increased their ability to answer natural language questions in recent years, due in large part to large pre-trained neural network models called Language Models. These language models have felled many benchmarks, including recently achieving an A grade on answering standardized multiple choice elementary science exams. But how much do these language models truly know about elementary science, and how robust is their knowledge? In this work, we present ScienceWorld, a new benchmark to test agents' scientific reasoning abilities. ScienceWorld is an interactive text game environment that tasks agents with performing 30 tasks drawn from the elementary science curriculum, like melting ice, building simple electrical circuits, using pollinators to help grow fruits, or understanding dominant versus recessive genetic traits. We show that current state-of-the-art language models that can easily answer elementary science questions, such as whether a metal fork is conductive or not, struggle when tasked to conduct an experiment to test this in a grounded, interactive environment, even with substantial training data. This presents the question of whether current models are simply retrieving answers to questions by way of observing a large number of similar input examples, or if they have learned to reason about concepts in a reusable manner. We hypothesize that agents need to be grounded in interactive environments to achieve such reasoning abilities. Our experiments provide empirical evidence supporting this hypothesis – showing that a 1.5 million parameter agent trained interactively for 100k steps outperforms an 11 billion parameter model statically trained for scientific question answering and reasoning via millions of expert demonstrations.

Bio: Peter Jansen is an Assistant Professor in the School of Information at the University of Arizona. A central focus of his research is how we can teach models to perform question answering while generating detailed human-readable explanations for their answers. His work is largely centered in the science domain, where he works to answer and explain standardized elementary and middle school science exams, as written, using a variety of automated inference formalisms. <http://www.cognitiveai.org/explanationbank/>

Keynote Talk: Text Toys and Glitch Poetics

Lynn Cherny
GhostWeather

Abstract: Not just a bug, the glitch offers creative possibility – especially in AI systems where we are travelers in a foggy latent space. The glitch is usually a visual metaphor, but it is alive and well in text encodings too. I’ll talk about projects (mine and others’) that explore neural spaces in poetic and game-like ways. Focusing on text play in this talk, we’ll visit media collages, mistaken translations, cross-modal cutups, and the dusty bottoms of game databases in search of the uncanny glitch that make us laugh because it’s true.

Bio: Lynn Cherny has a Ph.D. in linguistics from Stanford and worked in research on text-based virtual worlds (MUDs) and chat interfaces. She’s been in industry as a UX designer and data scientist for a hundred years now, at companies as diverse as Adobe, The Mathworks, TiVo, and Autodesk. She worked on the poetry generation algorithm for the [UK Dubai 2020 pavilion](#) and currently consults at Google Arts & Culture as an ML artist in residence, where she released a small game illustrated by [AI-generated images](#). Lynn has given keynotes and talks on creative AI projects at EuroPython and PyData conferences, Eyeo Festival, and KIKK Festival. [Her newsletter](#) is a fun mashup of weird folklore, AI art, NLP, games & narrative links.

Table of Contents

<i>A Systematic Survey of Text Worlds as Embodied Natural Language Environments</i> Peter Jansen	1
<i>A Minimal Computational Improviser Based on Oral Thought</i> Nick Montfort and Sebastian Bartlett Fernandez	16
<i>Craft an Iron Sword: Dynamically Generating Interactive Game Characters by Prompting Large Language Models Tuned on Code</i> Ryan Volum, Sudha Rao, Michael Xu, Gabriel DesGarences, Chris Brockett, Benjamin Van Durme, Olivia Deng, Akanksha Malhotra and Bill Dolan	25
<i>A Sequence Modelling Approach to Question Answering in Text-Based Games</i> Gregory Furman, Edan Toledo, Jonathan Phillip Shock and Jan Buys	44
<i>Automatic Exploration of Textual Environments with Language-Conditioned Autotelic Agents</i> Laetitia Teodorescu, Xingdi Yuan, Marc-Alexandre Côté and Pierre-Yves Oudeyer	59